

GAATTCTGGAG	GAATTATTCA	AAACATAAAC	ACAATAAAC	ATTTGAGTAG	TTGCCGCACA	60		
CACACACACA	CACAGCCCCT	GGATTATTAC	ACTAAAAGCG	ACACTCAA	TC	CAAAAAATCA	120	
GCAACAAAAA	CATCAATAAA	C	ATG CAT	TGG ATT	AAA TGT	TTA TTA	ACA GCA	171
			Met His	Trp Ile	Lys Cys	Leu Leu	Thr Ala	
		1		5		10		
TTC ATT TGC	TTC ACA	GTC ATC	GTG CAG	GTT CAC	AGT TCC	GGC AGC	TTT	219
Phe Ile	Cys Phe	Thr Val	Ile Val	Gln Val	His Ser	Ser Gly	Ser Phe	
15		20		25				
GAG TTG CGC	CTG AAG TAC	TTC AGC AAC	GAT CAC	GGG CGG	GAC AAC	GAG		267
Glu Leu Arg	Leu Lys Tyr	Phe Ser Asn	Asp His	Gly Arg	Asp Asn	Glu		
30		35		40				
GGT CGC TGC	TGC AGC GGG	GAG TCG GAC	GGA GCG	ACG GGC	AAG TGC	CTG		315
Gly Arg Cys	Cys Ser Gly	Glu Ser Asp	Gly Ala	Thr Gly	Lys Cys	Leu		
45		50		55				
GGC AGC TGC	AAG ACG CGG	TTT CGC GTC	TGC CTA	AAG CAC	TAC CAG	GCC		363
Gly Ser Cys	Lys Thr Arg	Phe Arg Val	Cys Leu	Lys His	Tyr Gln	Ala		
60		65		70				
ACC ATC GAC	ACC ACC TCC	CAG TGC ACC	TAC GGG GAC	GTG ATC ACG	CCC			411
Thr Ile Asp	Thr Ser Gln	Cys Thr Tyr	Gly Asp Val	Ile Thr Pro				
75		80		85		90		
ATT CTC GGC	GAG AAC TCG	GTC AAT CTG	ACC GAC	GCC CAG	CGC TTC	CAG		459
Ile Leu Gly	Glu Asn Ser	Val Asn Leu	Thr Asp Ala	Gln Arg	Phe Gln			
95		100		105				
AAC AAG GGC	TTC ACG AAT	CCC ATC CAG	TTC CCC	TTC TCG	TTC TCA	TGG		507
Asn Lys Gly	Phe Thr Asn	Pro Ile Gln	Phe Pro	Phe Ser	Phe Ser	Trp		
110		115		120				

CCG GGT ACC TTC TCG CTG ATC GTC GAG GCC TGG CAT GAT ACG AAC AAT Pro Gly Thr Phe Ser Leu Ile Val Glu Ala Trp His Asp Thr Asn Asn 125 130 135	555
AGC GGC AAT GCG CGA ACC AAC AAG CTC CTC ATC CAG CGA CTC TTG GTG Ser Gly Asn Ala Arg Thr Asn Lys Leu Leu Ile Gln Arg Leu Leu Val 140 145 150	603
CAG CAG GTA CTG GAG GTG TCC TCC GAA TGG AAG ACG AAC AAG TCG GAA Gln Gln Val Leu Glu Val Ser Ser Glu Trp Lys Thr Asn Lys Ser Glu 155 160 165 170	651
TCG CAG TAC ACG TCG CTG GAG TAC GAT TTC CGT GTC ACC TGC GAT CTC Ser Gln Tyr Thr Ser Leu Glu Tyr Asp Phe Arg Val Thr Cys Asp Leu 175 180 185	699
AAC TAC TAC GGA TCC GGC TGT GCC AAG TTC TGC CGG CCC CGC GAC GAT Asn Tyr Tyr Gly Ser Gly Cys Ala Lys Phe Cys Arg Pro Arg Asp Asp 190 195 200	747
TCA TTT GGA CAC TCG ACT TGC TCG GAG ACG GGC GAA ATT ATC TGT TTG Ser Phe Gly His Ser Thr Cys Ser Glu Thr Gly Glu Ile Ile Cys Leu 205 210 215	795
ACC GGA TGG CAG GGC GAT TAC TGT CAC ATA CCC AAA TGC GCC AAA GGC Thr Gly Trp Gln Gly Asp Tyr Cys His Ile Pro Lys Cys Ala Lys Gly 220 225 230	843
TGT GAA CAT GGA CAT TGC GAC AAA CCC AAT CAA TGC GTT TGC CAA CTG Cys Glu His Gly His Cys Asp Lys Pro Asn Gln Cys Val Cys Gln Leu 235 240 245 250	891
GGC TGG AAG GGA GCC TTG TGC AAC GAG TGC GTT CTG GAA CCG AAC TGC Gly Trp Lys Gly Ala Leu Cys Asn Glu Cys Val Leu Glu Pro Asn Cys 255 260 265	939

FIG.1B

ATC CAT GGC ACC TGC AAC AAA CCC TGG ACT TGC ATC TGC AAC GAG GGT Ile His Gly Thr Cys Asn Lys Pro Trp Thr Cys Ile Cys Asn Glu Gly 270 275 280	987
TGG GGA GGC TTG TAC TGC AAC CAG GAT CTG AAC TAC TGC ACC AAC CAC Trp Gly Gly Leu Tyr Cys Asn Gln Asp Leu Asn Tyr Cys Thr Asn His 285 290 295	1035
AGA CCC TGC AAG AAT GGC GGA ACC TGC TTC AAC ACC GGC GAG GGA TTG Arg Pro Cys Lys Asn Gly Gly Thr Cys Phe Asn Thr Gly Glu Gly Leu 300 305 310	1083
TAC ACA TGC AAA TGC GCT CCA GGA TAC AGT GGT GAT GAT TGC GAA AAT Tyr Thr Cys Lys Cys Ala Pro Gly Tyr Ser Gly Asp Asp Cys Glu Asn 315 320 325 330	1131
GAG ATC TAC TCC TGC GAT GCC GAT GTC AAT CCC TGC CAG AAT GGT GGT Glu Ile Tyr Ser Cys Asp Ala Asp Val Asn Pro Cys Gln Asn Gly Gly 335 340 345	1179
ACC TGC ATC GAT GAG CCG CAC ACA AAA ACC GGC TAC AAG TGT CAT TGC Thr Cys Ile Asp Glu Pro His Thr Lys Thr Gly Tyr Lys Cys His Cys 350 355 360	1227
GCC AAC GGC TGG AGC GGA AAG ATG TGC GAG GAG AAA GTG CTC ACG TGT Ala Asn Gly Trp Ser Gly Lys Met Cys Glu Glu Lys Val Leu Thr Cys 365 370 375	1275
TCG GAC AAA CCC TGT CAT CAG GGA ATC TGC CGC AAC GTT CGT CCT GGC Ser Asp Lys Pro Cys His Gln Gly Ile Cys Arg Asn Val Arg Pro Gly 380 385 390	1323
TTG GGA AGC AAG GGT CAG GGC TAC CAG TGC GAA TGT CCC ATT GGC TAC Leu Gly Ser Lys Gly Gln Gly Tyr Gln Cys Glu Cys Pro Ile Gly Tyr 395 400 405 410	1371

FIG.1C

AGC GGA CCC AAC TGC GAT CTC CAG CTG GAC AAC TGC AGT CCG AAT CCA Ser Gly Pro Asn Cys Asp Leu Gln Leu Asp Asn Cys Ser Pro Asn Pro 415 420 425	1419
TGC ATA AAC GGT GGA AGC TGT CAG CCG AGC GGA AAG TGT ATT TGC CCA Cys Ile Asn Gly Gly Ser Cys Gln Pro Ser Gly Lys Cys Ile Cys Pro 430 435 440	1467
GCG GGA TTT TCG GGA ACG AGA TGC GAG ACC AAC ATT GAC GAT TGT CTT Ala Gly Phe Ser Gly Thr Arg Cys Glu Thr Asn Ile Asp Asp Cys Leu 445 450 455	1515
GGC CAC CAG TGC GAG AAC GGA GGC ACC TGC ATA GAT ATG GTC AAC CAA Gly His Gln Cys Glu Asn Gly Gly Thr Cys Ile Asp Met Val Asn Gln 460 465 470	1563
TAT CGC TGC CAA TGC GTT CCC GGT TTC CAT GGC ACC CAC TGT AGT AGC Tyr Arg Cys Gln Cys Val Pro Gly Phe His Gly Thr His Cys Ser Ser 475 480 485 490	1611
AAA GTT GAC TTG TGC CTC ATC AGA CCG TGT GCC AAT GGA GGA ACC TGC Lys Val Asp Leu Cys Leu Ile Arg Pro Cys Ala Asn Gly Gly Thr Cys 495 500 505	1659
TTG AAT CTC AAC AAC GAT TAC CAG TGC ACC TGT CGT GCG GGA TTT ACT Leu Asn Leu Asn Asn Asp Tyr Gln Cys Thr Cys Arg Ala Gly Phe Thr 510 515 520	1707
GGC AAG GAT TGC TCT GTG GAC ATC GAT GAG TGC AGC AGT GGA CCC TGT Gly Lys Asp Cys Ser Val Asp Ile Asp Glu Cys Ser Ser Gly Pro Cys 525 530 535	1755
CAT AAC GGC GGC ACT TGC ATG AAC CGC GTC AAT TCG TTC GAA TGC GTG His Asn Gly Gly Thr Cys Met Asn Arg Val Asn Ser Phe Glu Cys Val 540 545 550	1803

FIG. 1D

338.37 1990 1991			
TGT	GCC	AAT	GGT
TTC	AGG	GGC	AAG
555	560	565	570
Cys	Ala	Asn	Gly
			Phe
			Arg
			Gly
			Lys
			Gln
			Cys
			Asp
			Glu
			Glu
			Ser
			Tyr
			Asp
TCG	GTG	ACC	TTC
GAT	GCC	CAC	CAA
575	580	585	
Ser	Val	Thr	Phe
			Asp
			Ala
			His
			Gln
			Tyr
			Gly
			Ala
			Thr
			Thr
			Gln
			Ala
			Arg
GCC	GAT	GGT	TTG
590	595	600	
ACC	AAT	GCC	CAG
			GTA
			GTC
			CTA
			ATT
			GCT
			GTT
			TTC
			TCC
Ala	Asp	Gly	Leu
			Thr
			Asn
			Ala
			Gln
			Val
			Val
			Leu
			Ile
			Ala
			Val
			Phe
			Ser
GTT	GCG	ATG	CCT
605	610	615	
TTG	GTG	GCG	GTT
ATT	GCG	GCG	GTC
			GTC
			TTC
			TGC
Val	Ala	Met	Pro
			Leu
			Val
			Ala
			Val
			Ile
			Ala
			Ala
			Cys
			Val
			Val
			Phe
			Cys
ATG	AAG	CGC	AAG
620	625	630	
CGT	AAG	CGT	GCT
AAG	CAG	CAG	CAG
635	640	645	
CGT	GAA	GAA	GAA
AAT	GCG	GTG	GCC
			ACA
			ATG
			CAT
			CAC
			AAT
			GCC
Met	Lys	Arg	Lys
			Arg
			Lys
			Arg
			Arg
			Ala
			Gln
			Glu
			Lys
			Asp
			Asp
			Ala
			Glu
			Ala
AGG	AAG	CAG	AAC
655	660	665	
AAC	GAA	CAG	CAG
670	675	680	
GAA	CAG	AAT	GCG
685	690	695	
CAT	GTG	GTG	GCT
			TCA
			GCC
			TCT
			CTG
			GGC
			GGC
			AAA
			ACT
Ser	Gly	Val	Gly
			Val
			Ala
			Leu
			Ala
			Ser
			Ala
			Leu
			Gly
			Gly
			Lys
			Thr
GGC	AGC	AAC	AGC
695	700	705	
GGT	CTC	ACC	TTC
AAC	GAT	GGC	GGC
710	715	720	
GGC	GGC	AAC	CCG
725	730	735	
AAT	ATC	ATC	ATC
Gly	Ser	Asn	Ser
			Gly
			Leu
			Thr
			Phe
			Asp
			Gly
			Gly
			Asn
			Pro
			Asn
			Ile
			Ile
AAA	AAC	ACC	TGG
740	745	750	
GAC	AAG	TCG	GTC
755	760	765	
AAC	AAC	ATT	TGT
770	775	780	
TGT	GCC	TCA	GCA
785	790	795	
GCA	GCA		
795	800	805	
Lys	Asn	Thr	Trp
			Asp
			Lys
			Ser
			Val
			Asn
			Asn
			Ile
			Cys
			Ala
			Ser
			Ala
			Ala

FIG.1E

GCA GCG GCG GCG GCG GCA GCA GCG GCG GAC GAG TGT CTC ATG TAC GGC 2283
 Ala Ala Ala Ala Ala Ala Ala Ala Asp Glu Cys Leu Met Tyr Gly
 700 705 710

GGA TAT GTG GCC TCG GTG GCG GAT AAC AAC AAT GCC AAC TCA GAC TTT 2331
 Gly Tyr Val Ala Ser Val Ala Asp Asn Asn Asn Ala Asn Ser Asp Phe
 715 720 725 730

TGT GTG GCT CCG CTA CAA AGA GCC AAG TCG CAA AAG CAA CTC AAC ACC 2379
 Cys Val Ala Pro Leu Gln Arg Ala Lys Ser Gln Lys Gln Leu Asn Thr
 735 740 745

GAT CCC ACG CTC ATG CAC CGC GGT TCG CCG GCA GGC AGC TCA GCC AAG 2427
 Asp Pro Thr Leu Met His Arg Gly Ser Pro Ala Gly Ser Ser Ala Lys
 750 755 760

GGA GCG TCT GGC GGA GGA CCG GGA GCG GCG GAG GGC AAG AGG ATC TCT 2475
 Gly Ala Ser Gly Gly Pro Gly Ala Ala Glu Gly Lys Arg Ile Ser
 765 770 775

GTT TTA GGC GAG GGT TCC TAC TGT AGC CAG CGT TGG CCC TCG TTG GCG 2523
 Val Leu Gly Glu Gly Ser Tyr Cys Ser Gln Arg Trp Pro Ser Leu Ala
 780 785 790

GCG GCG GGA GTG GCC GGA GCG TGT TCA TCC CAG CTA ATG GCT GCA GCT 2571
 Ala Ala Gly Val Ala Gly Ala Cys Ser Ser Gln Leu Met Ala Ala Ala
 795 800 805 810

TCG GCA GCG GGC AGC GGA GCG GGG ACG GCG CAA CAG CAG CGA TCC GTG 2619
 Ser Ala Ala Gly Ser Gly Ala Gly Thr Ala Gln Gln Gln Arg Ser Val
 815 820 825

GTC TGC GGC ACT CCG CAT ATG TAACTCCAAA AATCCGAAG GGCTCCTGGT 2670
 Val Cys Gly Thr Pro His Met
 830

AAATCCGGAG AAATCCGCAT GGAGGAGCTG ACAGCACATA CACAAAGAAA AGACTGGGTT 2730
 GGGTTCAAAA TGTGAGAGAG ACGCCAAAAT GTTGTGTTG ATTGAAGCAG TTTAGTCGTC 2790
 ACGAAAAATG AAAAATCTGT AACAGGCATA ACTCGTAAAC TCCCTAAAAA ATTTGTATAG 2850
 TAATTAGCAA AGCTGTGACC CAGCCGTTTC GATCCCGAAT TC 2892

FIG.1F

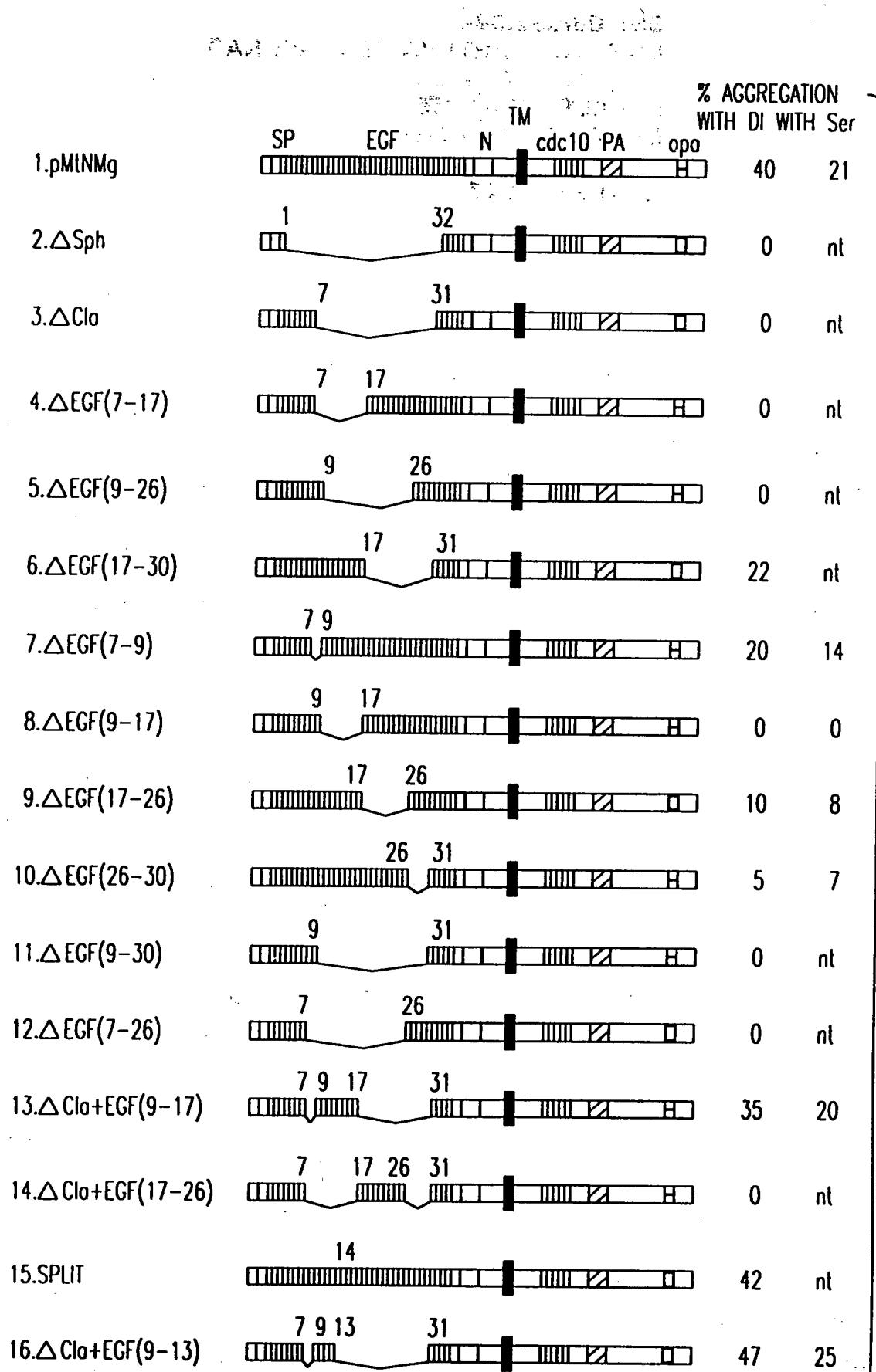


FIG.2A

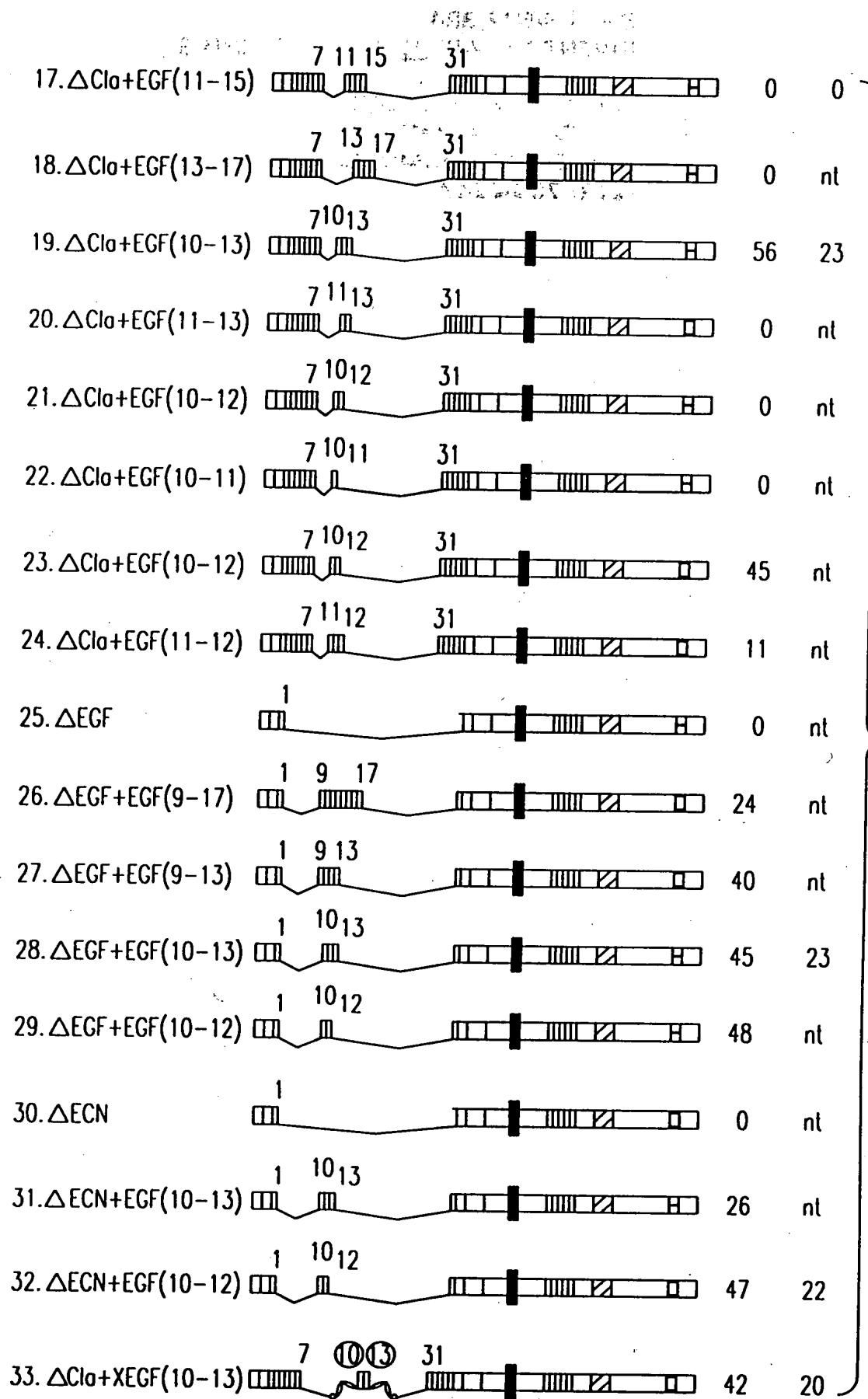


FIG.2B

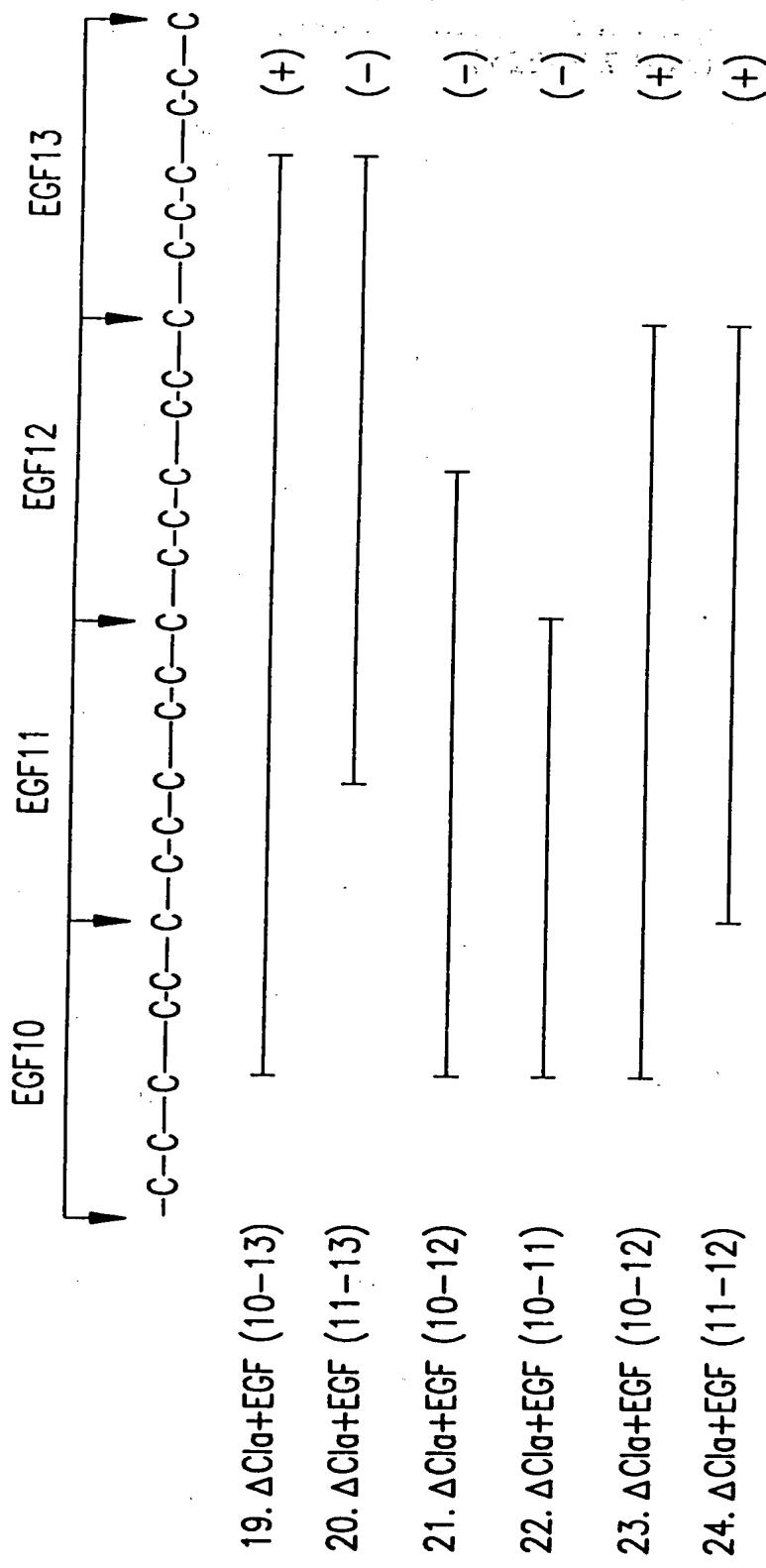


FIG. 3

	EGF 11	EGF 12
DROPHLLA	* * * * * EDIDECDQGS PCEHNGICVNTPGSYRCNCSSQDF	* * * * * TGPRCETNINECESHPCNEGSCLDDPGIFRCVCMPGFTGTQCE
XENOPUS	* NDVDECSLGANPCEHGGRCTNTLGSFGCNCPPQQYAGPRCEIDVNNECLS	* * * * * CNDSTCLDQIGEEFACICMPGYEGLYCE

FIG.4

1 CCGAGTCAGCCGGTGCCTGGCCCTTTCTGTCAACGGCTAAAGATC
121 AAGCACATACTAAGGTCATATAATAATAATAATTGTGTGATAACAACATTAT
241 GGCGTTATTTCAGCTATCCAGGCAAGTGTAGTGTGGAAAATAGAAACAAACAAAGGCA
361 CATACTCCAGAGTGAATTCGAAACAAACTCCATCTAGATCGCCAACCCAGGATCAGCGTCGGCA

481 TCGTCGGAGTCACAAATAGAACATAGCAGACAGCCTGGAAATGTCCAGAAAGACGGCG
SerSerLeuGluSerThrIleGluSerAlaAspSerLeuGlyMetSerLysLysThrAla

601 CGCGATTGTCGATCATTAAAGCTGGCTGCAACTTAATTGCTTAATTAAATACGTGTA-
ArgAspCysArgSerLeuLysSerAlaCysAsnLeuIleAlaLeuIleLeuLeu-

721 AACAGGCCATCTACTCAACGGCTATTGCTGCCAGGGCATGCCAGGGAACTTAGGGCCACCAAG
AsnSerHisLeuLeuAsnGlyTyrCysCysGlyMetProAlaGluLeuArgAlaThrLys

841 **ACCGAGCAGGGTGCAGGCATATCCACGGGCTGTTGGCAA** **GCACCAAGATA**
ThrGluGlnGlyIleSerIleSerThrGlyCysSerPheGlyAsnAlaThrThrLysIle **#2**

961 ACGGTTCTGGACGAAAGTCGTTACGGCTGATACTGCAGGGCTTGGATATGTCACACACAA
ThrPheArgTrpThrLysSerPheThrLeuIleLeuGlnAlaLeuAspMetTyrAsnThr
" "

1081 TCGCCGAACTGGACATCTGGACCAACATCGCCGGAATCACCTAACCTGTC
SerProGluIleTrpLysIleGlyArgAsnAlaGlyAlaIleThrTyrAlaGlyAl

11201 GACGATCAGTTGGTCACTACGGCTCCGAGGTCAAGCTCTGCCTGAATGGC
AspAspGlnPheGlyHisTyrAlaCysGlySerGluGlyGlnLysLeuCysLeuAsnGly

FIG. 5A

TACAAACATCAGGCCCTATCAAGTGGAAAGTGTCAAGTGTGAACAAACAAAAACGAGAG
 CCAAACAAACAAACAAACGAAAGGCAAAGTGGAAAAATGATACTGGAGAGTAC
 CCAAATCTGCATAACATGGCTAATTAAAGGCTGCCAGGGAAATTACATTGTGGTG
 AACGGCCCCAGAATGTACAAACATTAGGAAACATTTCGGGAAACACCAGCTACGTC
 MetPheArgGlySHisPheArgArgLysProAlaThrSer 13

ACAAAAAGGCAGCGTCCGGCATCGGGTACCCAAATGGCACCCCTGCCATGGAGGATC
 ThrLysArgGlnArgProArgHisArgValProLysLeuProSerThrIle 53

GTCCATAAGATAATCCGGCAGCTGGTAACTCGAGCTGGAAATTAGAAATCTCAAATACC
 ValHisLysIleSerAlaAlaGlyAsnPheGluLysLeuGluLysLeuAsnThr 93

#1

ACGATAGGCTGGCTGCCATGGCACCGACGGCATTCCGGCTGTGGCTGAAGGAGTACCAAGACC
 ThrIleGlyCysSerProCysThrThrAlaPheArgLeuCysLeuLysGluTyrGlnThr 133

CTGGGTGGCTCCAGGTTTGTGCTCAGCGATCCGGATTGGGAGGCCATTGTGCTGCCCTTT
 LeuGlyGlySerSerPheValLeuSerAspProGlyValGlyAlaIleValLeuProPhe 173

TCCTATCCAGATGGGAGAGGTTAATTGAGGAAACATCATACTCGGGCTGATACTGCCG
 SerTyrProAspAlaGluArgLeuValArgLeuIleGluGluThrSerTyrSerGlyValIleLeuPro 213

#4

CGGGTGGCAATGGGCCGTACCTACAAACACGGACCTGCACCGACCTTGCCGTCCGGCG
 ArgValGlyValCysAlaValThrTyrAsnThrThrCysThrCysArgProArg 253

TGGCAAGGGCGTCAACTGGGAGGGCCATATGCAAGGGGGCTGGCACCCGTCCACGGGC
 TrpGlnGlyValAsnCysGluGluValIleCysLysAlaGlyCysAspProValIHisGly 293

FIG. 5B

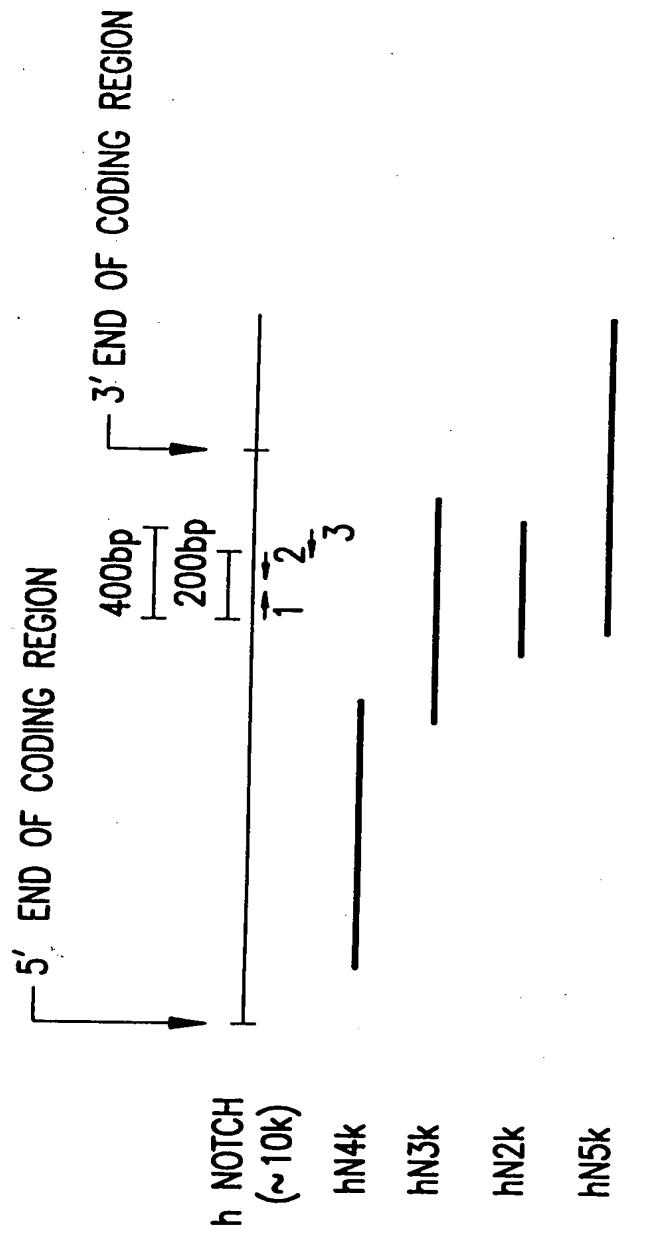
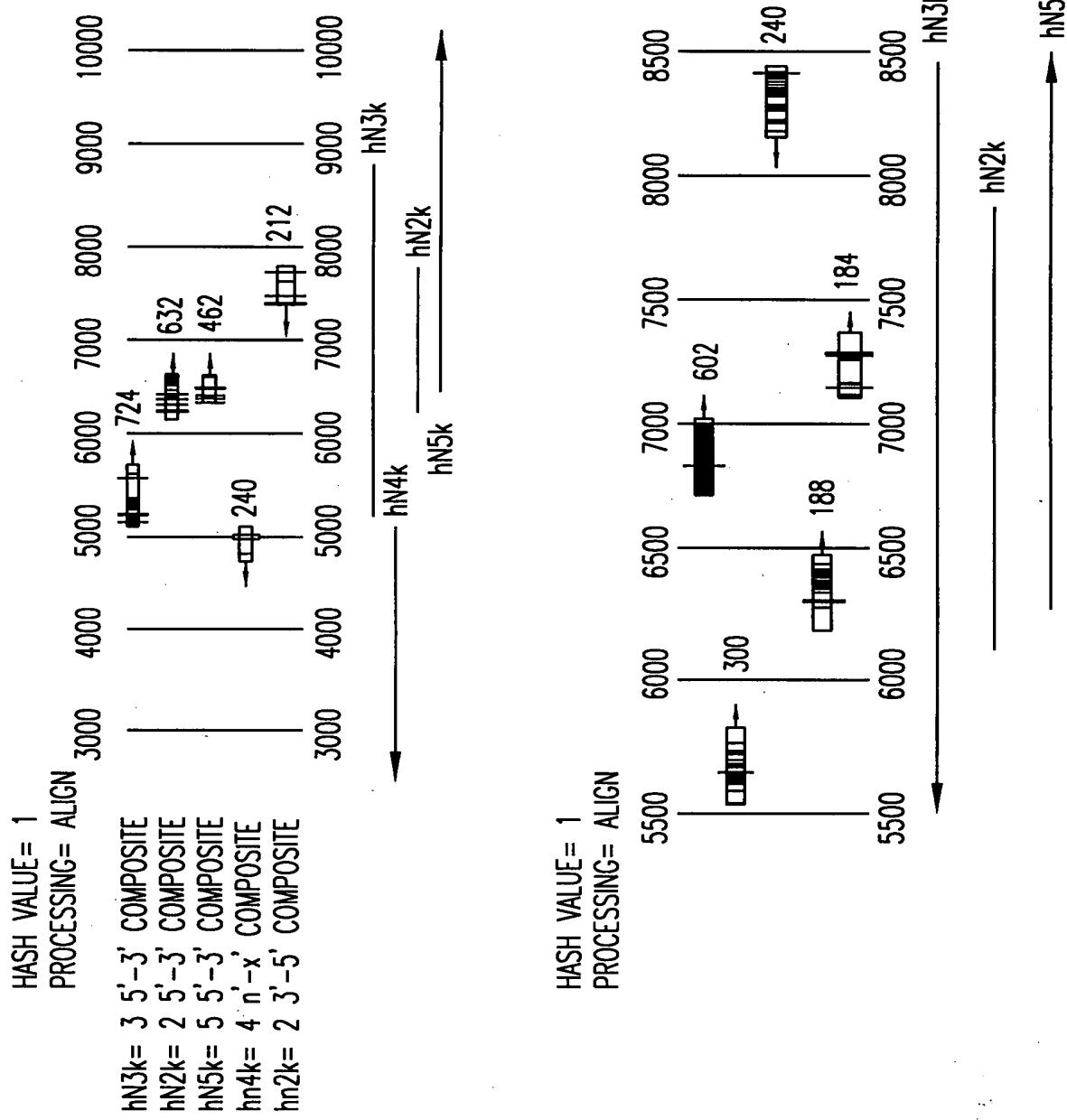


FIG. 6

FIG.7



1 GAATTCCGCT GGGAGAATGG TCTGAGCTAC CTGCCCGTCC TGCTGGGGCA TCAATGGCAA
61 GTGGGGAAAG CCACACTGGG CAAACGGGCC AGGCCATTTC TGGAATGTGG TACATGGTGG
121 GCAGGGGGCC CGCAACAGCT GGAGGGCAGG TGGACTGAGG CTGGGGATCC CCCGCTGGTT
181 GGGCAATACT GCCTTTACCC ATGAGCTGGA AAGTCACAAT GGGGGGCAAG GGCTCCCGAG
241 GGTGGTTATG TGCTTCCTTC AGGTGGC

FIG.8A

1 GAATTCCCTTC CATTATAACGT GACTTTCTG AAACTGTAGC CACCCCTAGTG TCTCTAACTC
61 CCTCTGGAGT TTGTCAGCTT TGGTCTTTTC AAAGAGCAGG CTCTCTTCAA GCTCCTTAAT
121 GCAGGGCATGC TCCAGTTGG TCTGCCTCTC AAGATCACCT TTGGTAATTG ATTCTTCTTC
181 AACCCGGAAC TGAAGGCTGG CTCTCACCCCT CTAGGCAGAG CAGGAATTCC GAGGTGGATG
241 TGTTAGATGT GAATGTCCGT GGCCCAGATG GCTGCACCCCC ATTGATGTTG GCTTCTCTCC
301 GAGGAGGGCAG CTCAGATTG AGTGTATGAAG ATGAAGATGC AGAGGACTGT TCTGCTAAC
361 TCATCACAGA CTTGGTCTAC CAGGGTGCCA GCCTCCAGAC CAGACAGACC GGACTGGTGA
421 GATGGCCCTG CACCTTGCAAG CCCGCTACTC ACGGGCTGAT GCTGCCAACG GTCTCCTGGA
481 TGCAGGTGCA GATGCCAATG CCCAGGACAA CATGGGCCGC TGTCCACTCC ATGCTGCAGT
541 GGCACGTGAT GCCAAGGTGT ATTCAAGATCT GTTA

FIG.8B

1 TCCAGATTCT GATTGCAAC CGAGTAAC TG ATCTAGATGC CAGGATGAAT GATGGTACTA
61 CACCCCTGAT CCTGGCTGCC CGCCTGGCTG TGGAGGGAAT GGTGGCAGAA CTGATCAACT
121 GCCAAGCGGA TGTGAATGCA GTGGATGACC ATGGAAAATC TGCTCTTCAC TGGGCAGCTG
181 CTGTCAATAA TGTGGAGGCA ACTCTTTGT TGTTGAAAAA TGGGGCCAAC CGAGACATGC
241 AGGACAACAA GGAAGAGACA CCTCTGTTTC TTGCTGCCGC GGAGGGAGCTA TAAGC

FIG.8C

1 GAATTCCATT CAGGAGGAAA GGGTGGGGAG AGAAGCAGGC ACCCACTTTC CCGTGGCTGG
61 ACTCGTTCCC AGGTGGCTCC ACCGGCAGCT GTGACCGCCG CAGGTGGGGG CGGAGTGCCA
121 TTCAGAAAAT TCCAGAAAAG CCCTACCCCA ACTCAGACGG CAACGTACA CCCGTGGGTA
181 GCAACTGGCA CACAAACAGC CAGCGTGTCT GGGGCACGGG GGGATGGCAC CCCCTGCAGG
241 CAGAGCTG

FIG.9A

1 CTAAAGGGAA CAAAAGCNNG AGCTCCACCG CGGGCGGCNC NGCTCTAGAA CTAGTGGANN
61 NCCCCGGGCTG CAGGAATTCC GGCGGACTGG GCTCGGGCTC AGAGCGGCAGC TGTGGAAGAG
121 ATTCTAGACC GGGAGAACAA GCGAATGGCT GACAGCTGGC CTCCAAAGTC ACCAGGCTCA
181 AATCGCTCGC CCTGGACATC GAGGGATGCA GAGGATCAGA ACCGGTACCT GGATGGCATG
241 ACTCGGATT ACAAGCATGA CCAGCCTGCT TACAGGGAGC GTGANNTTT CACATGCAGT
301 CGACAGACAC GAGCTCTATG CAT

FIG.9B

FIG. 10A

200	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	240
G	GC	TTT	GAC	TGC	CAG	CGT	GCG	GAA	GGC	CAG	TGC	AAC	CCC	CTG	TAC	GAC			
G	E	F	D	C	Q	R	A	E	G	Q	C	N	P	L	Y	D	>-		
210	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	240	
220	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	240	
230	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	240	
240	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	240	
250	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	240	
260	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	240	
270	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	240	
280	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	240	
290	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	240	
300	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	240	
310	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	240	
320	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	240	
330	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	240	
340	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	240	
350	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	240	
360	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	240	
370	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	240	
380	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	240	

FIG.10B

* 390 * * 400 * * 410 * * 420 * * 430 * *
 CCG GAG CAG CTG CGC AAC AGC TCC TTC CAC TTC CTG CGG GAG CTC AGC
 P E Q L R N S S F H E L R E L S>

 * 440 * * 450 * * 460 * * 470 * * 480 * *
 CGC GTG CTG CAC ACC ACC AAC GTG GTC TTC AAG CGT GAC GCA CAC GGC CAG
 R V L H T N V V F K R D A H G Q>

 490 * * 500 * * 510 * * 520 * *
 CAG ATG ATC TTC CCC TAC TAC GCC CCC GAG GAG GAG GAG CCC AAG GAC
 Q M I F P Y Y G R E E L R K H>

 530 * 540 * 550 * 560 * 570 *
 CCC ATC AAG CGT GCC GCC GAG GGC TGG GCA CCT GAC GCC CTG CTG
 P I K R A A E G W A A P D A L L>

FIG. 10C

580	*	*	590	*	*	600	*	*	610	*	*	620	*	
G	Q	V	K	A	S	L	L	P	G	G	S	E	G	G
630	*	*	640	*	*	650	*	*	660	*	*	670	*	
R	R	R	E	L	D	P	M	D	V	R	G	I	V	
680	*	*	690	*	*	700	*	*	710	*	*	720	*	
L	E	I	D	N	R	Q	C	V	Q	A	S	S	Q	
730	*	*	740	*	*	750	*	*	760	*	*	770	*	
Q	S	A	T	D	V	A	A	F	L	G	A	L	A	
											S	S	L>	

FIG. 10D

770	*	*	780	*	*	790	*	*	800	*	*	810	*	*	
GGC	AGC	CTC	AAC	ATC	CCC	TAC	AAG	ATC	GAG	GCC	GTG	CAG	AGT	GAG	ACC
G	S	L	N	T	P	Y	K	I	E	A	V	Q	S	E	T>-
820	*	*	830	*	*	840	*	*	850	*	*	860	*	*	
GTG	GAG	CCC	CCC	CCG	CCG	GCG	CAG	CTG	CAC	TTC	ATG	TAC	GTG	GCG	GCG
V	E	P	P	P	A	Q	L	H	F	M	Y	V	A	A>	
870	*	*	880	*	*	890	*	*	900	*	*	910	*	*	
GCC	GCC	TTR	GTG	CTT	CTG	TTC	TTC	GTG	GGC	TGC	GGG	GTG	CTG	CTG	TCC
A	A	F	V	L	L	F	F	V	G	C	G	V	L	L	S>
920	*	*	930	*	*	940	*	*	950	*	*	960	*	*	
CGC	AAG	CGC	CGG	CGG	CAG	CAT	GGC	CAG	CTC	TGG	TTC	CCT	GAG	GGC	TTC
R	K	R	R	R	Q	H	G	Q	L	W	F	P	E	G	F>

FIG.10E

	970	980	990	1000
	*	*	*	*
AAA	GTG TCT GAG GCC AGC AAG AAG	CGG CGG GAG CCC CTC GGC GAC		
K	V S E A	K K R R E	P L	G E>
	1010	1020	1030	1040
	*	*	*	*
GAC	TCC GTG GGC CTC AAG CCC CTG AAC GCT TCA GAC GGT GCC CTC			*
D	S V G L K P	N A S D G A L>		
	1060	1070	1080	1090
	*	*	*	*
ATG	GAC AAC CAG AAT GAG TGG GGG GAC GAG GAC CTG GAG ACC AAG			*
M	D D N Q N E W G D E D L E T K>			
	1110	1120	1130	1140
	*	*	*	*
AAG	TTC CGG TTC GAG GAG CCC GTG GTT CTG CCT GAC CTG GAC GAC CAG			*
K	F R E P V V L P D L D D Q>			

FIG.10F

1160	*	1170	*	1180	*	1190	*	1200	*					
T	D	H	R	Q	W	T	Q	H	L	D	A	A	D	L>
*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
ACA	GAC	CAC	CGC	CAG	TGG	ACT	CAG	CAC	CTG	GAT	GCC	GCT	GAC	CTG

1210	*	1220	*	1230	*	1240	*	1250	*	1260	*	1270	*	1280	*	1290	*
R	M	S	A	M	A	P	T	P	Q	G	E	V	D	A>	G	GCC	GAC
*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
GCG	ATG	TCT	GCC	ATG	GCC	CCC	ACA	CCC	CAG	GGT	GAG	GTT	GAC	GCC			

1250	*	1260	*	1270	*	1280	*	1290	*	1300	*	1310	*	1320	*	1330	*	1340	*
D	C	M	D	V	N	V	R	G	P	D	G	F	T	P	D>	G	CTC	CCC	
*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*		
GAC	TGC	ATG	GAC	GTC	AAT	GTC	CGC	GGG	CCT	GAT	GGC	TTC	ACC	CCC	CTC				

1300	*	1310	*	1320	*	1330	*	1340	*	1350	*	1360	*	1370	*	1380	*	1390	*
M	I	A	S	C	S	G	G	G	L	E	T	G	N	S	E>	G	GAG	GAC	
*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*		
ATG	ATC	GGC	TCC	TGC	AGC	GGC	GGC	GGC	CTG	GAG	ACG	GGC	AAC	AGC	GAG				

F16.106

1350	*	*	1360	*	*	1370	*	*	1380	*	*	1390	*					
*	*	*	GAA	GAG	GAC	GCG	CCG	GCC	GTC	ATC	TCC	GAC	TTC	ATC	TAC	CAG	GGC	
E	E	E	D	A	P	A	V	I	S	D	F	I	Y	Q	G>			
1400	*	*	1410	*	*	1420	*	*	1430	*	*	1440	*	*	*	*	*	
*	*	*	GCC	AGC	CTG	CAC	AAC	CAG	ACA	GAC	CCG	ACG	GGC	GAG	ACC	GCC	TTG	CAC
A	S	L	H	N	Q	T	D	R	T	G	E	T	A	L	H>			
1450	*	*	1460	*	*	1470	*	*	1480	*	*	1490	*	*	1500	*	*	
*	*	*	CTG	GCC	GCC	TAC	TCA	CGC	TCT	GAT	GCC	AAG	CGC	CTG	CTG	GAG		
L	A	A	R	Y	S	R	S	D	A	A	K	R	L	L	E>			
1510	*	*	1520	*	*	1530	*	*										
GCC	AGC	GCA	GAT	GCC	AAC	ATC	CAG	GAC	AAC	ATG	GGC	CGC	ACC	CCG	CTG			
A	S	A	D	A	N	I	Q	D	N	M	G	R	T	P	L>			

FIG. 10H

1540	*	*	1550	*	*	1560	*	*	1570	*	*	1580	*		
CAT GCG GCT GTG TCT GCC GAC GCA CAA GGT GTC TTC CAG ATC CTG ATC															
H	A	A	V	S	A	D	A	Q	G	V	F	Q	I	L	I>
1590	*	*	1600	*	*	1610	*	*	1620	*	*	1630	*		
CGG AAC CGA GCC ACA GCA CTG GAT GCC CGC ATG CAT GAT GGC ACC ACG															
R	N	R	A	T	D	L	D	A	R	M	H	D	G	T	T>
1640	*	*	1650	*	*	1660	*	*	1670	*	*	1680	*		
CCA CTG ATC CTG GCT GCC CGC CTG GCC GTG GAG GGC ATG CTG GAG GAC															
P	L	I	L	A	A	R	L	A	V	E	G	M	L	E	D>
1690	*	*	1700	*	*	1710	*	*	1720	*	*	1730	*		
CTC ATC AAC TCA CAC GCC GAC AAC GCC GTA GAT GAC CTG GGC AAG															
L	I	N	S	H	A	D	V	N	A	V	D	D	L	G	K>
1740	*	*	1750	*	*	1760	*	*	1770	*	*	1780	*		
TCC GCC CTG CAC TGG GCC CCC GTG AAC AAT GTC GAT GCC GCA GTT															
S	A	L	H	A	A	A	V	N	N	V	D	A	A	V>	

FIG. 101

1780	*	*	1790	*	*	1800	*	*	1810	*	*	1820	*				
	GTG	CTC	CTG	AAG	AAC	GGG	GCT	AAC	AAA	GAT	ATG	CAG	AAC	AGG	GAG		
V	L	L	K	N	G	A	A	N	K	D	M	Q	N	N	R	E>	
1830	*	*	1840	*	*	1850	*	*	1860	*	*	1870	*	*			
	GAG	ACA	CCC	CTC	CTG	TTT	CTG	GCC	GCC	GGG	GAG	GGC	AGC	TAC	GAG	ACC	GCC
E	T	P	L	F	L	A	A	R	E	G	S	Y	E	T	A>		
1880	*	*	1890	*	*	1900	*	*	1910	*	*	1920	*	*			
	AAG	GTG	CTG	CTG	GAC	CAC	TTT	GCC	AAC	CGG	GAC	ATC	ACG	GAT	CAT	ATG	
K	V	L	L	D	H	F	A	N	R	D	I	T	D	H	M>		
1930	*	*	1940	*	*	1950	*	*	1960	*	*	1970	*	*			
	GAC	CGC	CTG	CCG	CGC	GAC	ATC	GCA	CAG	GAG	CGC	ATG	CAT	CAC	GAC	ATC	
D	R	L	P	R	D	I	A	Q	E	R	M	H	H	D	I>		

FIG.10J

1970	*	*	*	*	*	*	*	*	*	*	*
	GTG	AGG	CTG	CTG	GAC	GAG	TAC	AAC	CTG	GTG	CGC
V	R	L	L	D	E	Y	N	L	V	R	S
											H>
1980	*	*	*	*	*	*	*	*	*	*	*
	GTG	AGG	CTG	CTG	GAC	GAG	TAC	AAC	CTG	GTG	CGC
V	R	L	L	D	E	Y	N	L	V	R	S
											H>
1990	*	*	*	*	*	*	*	*	*	*	*
	GTG	AGG	CTG	CTG	GAC	GAG	TAC	AAC	CTG	GTG	CGC
V	R	L	L	D	E	Y	N	L	V	R	S
											H>
2000	*	*	*	*	*	*	*	*	*	*	*
	GTG	AGG	CTG	CTG	GAC	GAG	TAC	AAC	CTG	GTG	CGC
V	R	L	L	D	E	Y	N	L	V	R	S
											H>
2010	*	*	*	*	*	*	*	*	*	*	*
	GTG	AGG	CTG	CTG	GAC	GAG	TAC	AAC	CTG	GTG	CGC
V	R	L	L	D	E	Y	N	L	V	R	S
											H>
2020	*	*	*	*	*	*	*	*	*	*	*
	GGG										
G	A	P	L	G	G	T	P	T	L	S	P
											H>
2030	*	*	*	*	*	*	*	*	*	*	*
	GGG										
G	A	P	L	G	G	T	P	T	L	S	P
											H>
2040	*	*	*	*	*	*	*	*	*	*	*
	GGG										
G	A	P	L	G	G	T	P	T	L	S	P
											H>
2050	*	*	*	*	*	*	*	*	*	*	*
	GGG										
G	A	P	L	G	G	T	P	T	L	S	P
											H>
2060	*	*	*	*	*	*	*	*	*	*	*
	GGG										
G	A	P	L	G	G	T	P	T	L	S	P
											H>
2070	*	*	*	*	*	*	*	*	*	*	*
	CCC	AAC	GGC	TAC	CTG	GGC	AGC	CTC	AAC	CCC	GGC
P	N	G	Y	L	G	S	L	K	P	G	V
											H>
2080	*	*	*	*	*	*	*	*	*	*	*
	CCC	AAC	GGC	TAC	CTG	GGC	AGC	CTC	AAC	CCC	GGC
P	N	G	Y	L	G	S	L	K	P	G	V
											H>
2090	*	*	*	*	*	*	*	*	*	*	*
	CCC	AAC	GGC	TAC	CTG	GGC	AGC	CTC	AAC	CCC	GGC
P	N	G	Y	L	G	S	L	K	P	G	V
											H>
2100	*	*	*	*	*	*	*	*	*	*	*
	CCC	AAC	GGC	TAC	CTG	GGC	AGC	CTC	AAC	CCC	GGC
P	N	G	Y	L	G	S	L	K	P	G	V
											H>
2110	*	*	*	*	*	*	*	*	*	*	*
	CCC	AAC	GGC	TAC	CTG	GGC	AGC	CTC	AAC	CCC	GGC
P	N	G	Y	L	G	S	L	K	P	G	V
											H>
2120	*	*	*	*	*	*	*	*	*	*	*
	GTC	CCG	AAG	CCC	AGC	AGC	AAA	GGC	CTG	GGC	TGT
V	R	K	P	S	S	K	G	L	A	C	G
											H>
2130	*	*	*	*	*	*	*	*	*	*	*
	GTC	CCG	AAG	CCC	AGC	AGC	AAA	GGC	CTG	GGC	AAG
V	R	K	P	S	S	K	G	L	A	C	G
											H>
2140	*	*	*	*	*	*	*	*	*	*	*
	GTC	CCG	AAG	CCC	AGC	AGC	AAA	GGC	CTG	GGC	AAG
V	R	K	P	S	S	K	G	L	A	C	G
											H>
2150	*	*	*	*	*	*	*	*	*	*	*
	GTC	CCG	AAG	CCC	AGC	AGC	AAA	GGC	CTG	GGC	AAG
V	R	K	P	S	S	K	G	L	A	C	G
											H>
2160	*	*	*	*	*	*	*	*	*	*	*
	GTC	CCG	AAG	CCC	AGC	AGC	AAA	GGC	CTG	GGC	AAG
V	R	K	P	S	S	K	G	L	A	C	G
											H>

FIG. 10K

2170	*	*	2180	*	*	2190	*	*	2200	*	*				
AAG	GAC	CTC	AAG	GCA	CGG	AGG	AAG	TCC	CAG	GAT	GGC	AAG	GGC	TGC	
K	D	L	K	A	R	R	K	S	Q	D	G	K	G	C>	
2210	*	*	2220	*	*	2230	*	*	2240	*	*	2250	*	*	
CTG	CTG	GAC	AGC	TCC	GGC	ATG	CTG	TCG	CCC	GTG	GAC	TCC	CTG	GAG	TCA
L	L	D	S	S	G	M	L	S	P	V	D	S	L	E	S>
2260	*	*	2270	*	*	2280	*	*	2290	*	*	2300	*	*	
CCC	CAT	GGC	TAC	CTG	TCA	GAC	GTG	GCC	TCG	CCG	CCA	CTG	CTG	CCC	TCC
P	H	G	Y	L	S	D	V	A	S	P	P	L	L	P	S>
2310	*	*	2320	*	*	2330	*	*	2340	*	*	2350	*	*	
CCG	TTC	CAG	CAG	TCT	CCG	TCC	GTG	CCC	CTG	CTG	CTG	CTG	CCT	GGG	ATG
P	F	Q	Q	S	P	S	V	P	N	H	L	P	G	ND	

FIG. 10L

2360	2370	2380	2390	2400
CCC GAC ACC CAC	CTG GGC ATC GGG CAC	CTG AAC GTC GCG GCC	AAG CCC	
P D T H	L G I G H	L N V A A	A K P >	
2410	2420	2430	2440	
GAG ATG GCG GCG	CTG GGT GGG GGC	CTG GCC TTT GAG ACT GGC		
E M A A	L G G G	L A F E T G		
2450	2460	2470	2480	2490
CCA CCTT CGT CTC	TCC CAC CCT GTG	GCC TCT GGC ACC ACC GTC		
P P R L	S H L P V A	S G T S T V >		
2500	2510	2520	2530	2540
CTG GGC TCC AGC AGC	GGA GGG GCC CTG AAT	TTC ACT GTG GGC GGG TCC		
L G S S	S G G A L N F	V T V G G S >		

FIG.10M

2550	*	*	2560	*	*	2570	*	*	2580	*	*	2590	*		
ACC	AGT	TTG	AAT	GGT	CAA	TGG	GAG	TGG	CTG	TCC	CGG	CTG	AGC	GGC	
T	S	L	N	G	Q	C	E	W	L	S	R	L	Q	S	G>
2600	*	*	2610	*	*	2620	*	*	2630	*	*	2640	*		
ATG	CTG	CCG	AAC	CAA	TAC	AAC	CCT	CTG	CGG	GGG	AGT	GTG	GCA	CCA	GGC
M	V	P	N	Q	Y	N	P	L	R	G	S	V	A	P	G>
2650	*	*	2660	*	*	2670	*	*	2680	*	*	2690	*		
CCC	CTG	AGC	ACA	CAG	GCC	CCC	TCC	CTG	CAG	CAT	GGC	ATG	GTA	GGC	CCG
P	L	S	T	Q	A	P	S	L	Q	H	G	M	V	G	P>
2700	*	*	2710	*	*	2720	*	*	2730	*	*	2740	*		
CTG	CAC	AGT	AGC	CTT	GCT	GCC	AGC	GCC	CTG	TCC	CAG	ATG	ATG	AGC	TAC
L	H	S	S	L	A	A	S	A	L	S	Q	M	M	S	Y>

FIG. 10N

2740	*	*	2750	*	*	2760	*	*	2770	*	*	2780	*		
CAG GGC CTG CCC AGC ACC CGG CTG GCC ACC CAG CCT CAC CTG GTG CAG															
Q	G	L	P	S	T	R	L	A	T	Q	P	H	L	V	Q>
2790	*	*	2800	*	*	2810	*	*	2820	*	*	2830	*		
ACC CAG CAG GTG CAG CCA CAA AAC TTA CAG ATG CAG CAG AAC CTC															
T	Q	Q	V	Q	P	Q	N	L	Q	M	Q	Q	N	L>	
2840	*	*	2850	*	*	2860	*	*	2870	*	*	2880	*		
CAG CCA GCA AAC ATC CAG CAG CAA AGC CTC CAG CCA CCA CCA															
Q	P	A	N	I	Q	Q	Q	S	L	Q	P	P	P	P>	
2890	*	*	2900	*	*	2910	*	*	2920	*	*	2880	*		
CCA CCA CAG CCG CAC CTT GCC GTG AGC TCA GCA GCC AGC GGC CAC CTC															
P	P	Q	P	H	L	G	V	S	S	A	A	S	G	H	L>
2930	*	*	2940	*	*	2950	*	*	2960	*	*	2970	*		
GGC CGG AGC TTC CTG AGT GGA GAG CCC AGC CAG GCA GAC GTG CAG CCA															
G	R	S	F	L	S	G	E	P	S	Q	A	D	V	Q	P>

FIG.100

2980	*	*	2990	*	*	3000	*	*	3010	*	*	3020	*			
L	G	C	CCC	GGC	AGC	CTG	GCG	GTG	CAC	ACT	ATT	CTG	CCC	CAG	GAG	AGC
P	S	S	L	A	V	H	T	I	L	P	Q	E	S	>		
3030	*	*	3040	*	*	3050	*	*	3060	*	*	3070	*	*		
P	A	L	CCC	CTG	CCC	ACG	TCG	CCA	TCC	TCG	CTG	GTC	CCA	CCC	GTG	ACC
A	L	P	T	S	L	P	S	L	P	S	L	V	P	P	V	T>
3080	*	*	3090	*	*	3100	*	*	3110	*	*	3120	*	*		
A	A	Q	CAG	TTC	CTG	ACG	CCC	CCC	TCC	CAG	CAC	AGC	TAC	TCC	TCG	CCT
		F	L	T	P	P	Q	S	Q	H	S	Y	S	S	P	>

FIG. 10P

3130	*	*	3140	*	*	3150	*	*	3160	*	*	
GTG	GAC	AAC	ACC	CCC	AGC	CAC	CAG	CTA	CAG	GTC	CCT	GTT
V	D	N	T	P	S	H	Q	L	Q	CCT	CCT	GTA
									V	P	P	ATG
										V	V	M>

3170	*	*	3180	*	*	3190	*	*	3200	*	*	3210	*	
GTA	ATG	ATC	CGA	TCT	TCG	GAT	CCT	TCT	AAA	GGC	TCA	TCA	ATT	TTG
V	M	I	R	S	S	D	P	S	K	G	S	S	I	I>

3220	*	*	3230	*	*
GAA	GCT	CCC	GAC	TCA	TGG
E	A	P	D	S	M>

FIG. 10Q

100 200 300 400 500 600 700 800 900 1000

G GAG GTG GAT GTG TTA GAT GTG AAT GTC CGT GGC CCA GAT GGC TGC 46
 Glu Val Asp Val Leu Asp Val Asn Val Arg Gly Pro Asp Gly Cys
 1 5 10 15

ACC CCA TTG ATG TTG GCT TCT CTC CGA GGA GGC AGC TCA GAT TTG AGT 94
 Thr Pro Leu Met Leu Ala Ser Leu Arg Gly Gly Ser Ser Asp Leu Ser
 20 25 50

GAT GAA GAT GAA GAT GCA GAG GAC TCT TCT GCT AAC ATC ATC ACA GAC 142
 Asp Glu Asp Glu Asp Ala Glu Asp Ser Ser Ala Asn Ile Ile Thr Asp
 35 40 45

TTG GTC TAC CAG GGT GCC AGC CTC CAG GCC CAG ACA GAC CGG ACT GGT 190
 Leu Val Tyr Gln Gly Ala Ser Leu Gln Ala Gln Thr Asp Arg Thr Gly
 50 55 60

GAG ATG GCC CTG CAC CTT GCA GCC CGC TAC TCA CGG GCT GAT GCT GCC 238
 Glu Met Ala Leu His Leu Ala Ala Arg Tyr Ser Arg Ala Asp Ala Ala
 65 70 75

AAG CGT CTC CTG GAT GCA GGT GCA GAT GCC AAT GCC CAG GAC AAC ATG 286
 Lys Arg Leu Leu Asp Ala Gly Ala Asp Ala Asn Ala Gln Asp Asn Met
 80 85 90 95

GGC CGC TGT CCA CTC CAT GCT GCA GTG GCA GCT GAT GCC CAA GGT GTC 334
 Gly Arg Cys Pro Leu His Ala Ala Val Ala Ala Asp Ala Gln Gly Val
 100 105 110

TTC CAG ATT CTG ATT CGC AAC CGA GTA ACT GAT CTA GAT GCC AGG ATG 382
 Phe Gln Ile Leu Ile Arg Asn Arg Val Thr Asp Leu Asp Ala Arg Met
 115 120 125

AAT GAT GGT ACT ACA CCC CTG ATC CTG GCT GCC CGC CTG GCT GTG GAG 430
 Asn Asp Gly Thr Thr Pro Leu Ile Leu Ala Ala Arg Leu Ala Val Glu
 130 135 140

FIG.11A

+ 1000 1000 1000

GGA ATG GTG GCA GAA CTG ATC AAC TGC CAA GCG GAT GTG AAT GCA GTG 478
 Gly Met Val Ala Glu Leu Ile Asn Cys Gln Ala Asp Val Asn Ala Val
 145 150 155

GAT GAC CAT GGA AAA TCT GCT CTT CAC TGG GCA GCT GCT GTC AAT AAT 526
 Asp Asp His Gly Lys Ser Ala Leu His Trp Ala Ala Val Asn Asn
 160 165 170 175

GTG GAG GCA ACT CTT TTG TTG AAA AAT GGG GCC AAC CGA GAC ATG 574
 Val Glu Ala Thr Leu Leu Leu Lys Asn Gly Ala Asn Arg Asp Met
 180 185 190

CAG GAC AAC AAG GAA GAG ACA CCT CTG TTT CTT GCT GCC CGG GAG GGG 622
 Gln Asp Asn Lys Glu Glu Thr Pro Leu Phe Leu Ala Ala Arg Glu Gly
 195 200 205

AGC TAT GAA GCA GCC AAG ATC CTG TTA GAC CAT TTT GCC AAT CGA GAC 670
 Ser Tyr Glu Ala Ala Lys Ile Leu Leu Asp His Phe Ala Asn Arg Asp
 210 215 220

ATC ACA GAC CAT ATG GAT CGT CTT CCC CGG GAT GTG GCT CGG GAT CGC 718
 Ile Thr Asp His Met Asp Arg Leu Pro Arg Asp Val Ala Arg Asp Arg
 225 230 235

ATG CAC CAT GAC ATT GTG CGC CTT CTG GAT GAA TAC AAT GTG ACC CCA 766
 Met His His Asp Ile Val Arg Leu Leu Asp Glu Tyr Asn Val Thr Pro
 240 245 250 255

AGC CCT CCA GGC ACC GTG TTG ACT TCT GCT CTC TCA CCT GTC ATC TGT 814
 Ser Pro Pro Gly Thr Val Leu Thr Ser Ala Leu Ser Pro Val Ile Cys
 260 265 270

GGG CCC AAC AGA TCT TTC CTC AGC CTG AAG CAC ACC CCA ATG GGC AAG 862
 Gly Pro Asn Arg Ser Phe Leu Ser Leu Lys His Thr Pro Met Gly Lys
 275 280 285

FIG.11B

AAG TCT AGA CGG CCC AGT GCC AAG AGT ACC ATG CCT ACT AGC CTC CCT 910
 Lys Ser Arg Arg Pro Ser Ala Lys Ser Thr Met Pro Thr Ser Leu Pro
 290 295 300

AAC CTT GCC AAG GAG GCA AAG GAT GCC AAG GGT AGT AGG AGG AAG AAG 958
 Asn Leu Ala Lys Glu Ala Lys Asp Ala Lys Gly Ser Arg Arg Lys Lys
 305 310 315

TCT CTG AGT GAG AAG GTC CAA CTG TCT GAG AGT TCA GTA ACT TTA TCC 1006
 Ser Leu Ser Glu Lys Val Gln Leu Ser Glu Ser Ser Val Thr Leu Ser
 320 325 330 335

CCT GTT GAT TCC CTA GAA TCT CCT CAC ACG TAT GTT TCC GAC ACC ACA 1054
 Pro Val Asp Ser Leu Glu Ser Pro His Thr Tyr Val Ser Asp Thr Thr
 340 345 350

TCC TCT CCA ATG ATT ACA TCC CCT GGG ATC TTA CAG GCC TCA CCC AAC 1102
 Ser Ser Pro Met Ile Thr Ser Pro Gly Ile Leu Gln Ala Ser Pro Asn
 355 360 365

CCT ATG TTG GCC ACT GCC CCT CCT GCC CCA GTC CAT GCC CAG CAT 1150
 Pro Met Leu Ala Thr Ala Ala Pro Pro Ala Pro Val His Ala Gln His
 370 375 380

GCA CTA TCT TTT TCT AAC CTT CAT GAA ATG CAG CCT TTG GCA CAT GGG 1198
 Ala Leu Ser Phe Ser Asn Leu His Glu Met Gln Pro Leu Ala His Gly
 385 390 395

GCC AGC ACT GTG CTT CCC TCA GTG AGC CAG TTG CTA TCC CAC CAC CAC 1246
 Ala Ser Thr Val Leu Pro Ser Val Ser Gln Leu Leu Ser His His His
 400 405 410 415

ATT GTG TCT CCA GGC AGT GGC AGT GCT GGA AGC TTG AGT AGG CTC CAT 1294
 Ile Val Ser Pro Gly Ser Gly Ser Ala Gly Ser Leu Ser Arg Leu His
 420 425 430

CCA GTC CCA GTC CCA GCA GAT TGG ATG AAC CGC ATG GAG GTG AAT GAG 1342
 Pro Val Pro Val Pro Ala Asp Trp Met Asn Arg Met Glu Val Asn Glu
 435 440 445

FIG.11C

ACC CAG TAC AAT GAG ATG TTT GGT ATG GTC CTG GCT CCA GCT GAG GGC 1390
 Thr Gln Tyr Asn Glu Met Phe Gly Met Val Leu Ala Pro Ala Glu Gly
 450 455 460

ACC CAT CCT GGC ATA GCT CCC CAG AGC AGG CCA CCT GAA GGG AAG CAC 1438
 Thr His Pro Gly Ile Ala Pro Gln Ser Arg Pro Pro Glu Gly Lys His
 465 470 475

ATA ACC ACC CCT CGG GAG CCC TTG CCC CCC ATT GTG ACT TTC CAG CTC 1486
 Ile Thr Thr Pro Arg Glu Pro Leu Pro Pro Ile Val Thr Phe Gln Leu
 480 485 490 495

ATC CCT AAA GGC AGT ATT GCC CAA CCA GCG GGG GCT CCC CAG CCT CAG 1534
 Ile Pro Lys Gly Ser Ile Ala Gln Pro Ala Gly Ala Pro Gln Pro Gln
 500 505 510

TCC ACC TGC CCT CCA GCT GTT GCG GGC CCC CTG CCC ACC ATG TAC CAG 1582
 Ser Thr Cys Pro Pro Ala Val Ala Gly Pro Leu Pro Thr Met Tyr Gln
 515 520 525

ATT CCA GAA ATG GCC CGT TTG CCC AGT GTG GCT TTC CCC ACT GCC ATG 1630
 Ile Pro Glu Met Ala Arg Leu Pro Ser Val Ala Phe Pro Thr Ala Met
 530 535 540

ATG CCC CAG CAG GAC GGG CAG GTA GCT CAG ACC ATT CTC CCA GCC TAT 1678
 Met Pro Gln Gln Asp Gly Gln Val Ala Gln Thr Ile Leu Pro Ala Tyr
 545 550 555

CAT CCT TTC CCA GCC TCT GTG GGC AAG TAC CCC ACA CCC CCT TCA CAG 1726
 His Pro Phe Pro Ala Ser Val Gly Lys Tyr Pro Thr Pro Pro Ser Gln
 560 565 570 575

CAC AGT TAT GCT TCC TCA AAT GCT GAG CGA ACA CCC AGT CAC AGT 1774
 His Ser Tyr Ala Ser Ser Asn Ala Ala Glu Arg Thr Pro Ser His Ser
 580 585 590

GGT CAC CTC CAG GGT GAG CAT CCC TAC CTG ACA CCA TCC CCA GAG TCT 1822
 Gly His Leu Gln Gly Glu His Pro Tyr Leu Thr Pro Ser Pro Glu Ser
 595 600 605

FIG.11D

CCT GAC CAG TGG TCA AGT TCA TCA CCC CAC TCT GCT TCT GAC TGG TCA Pro Asp Gln Trp Ser Ser Ser Pro His Ser Ala Ser Asp Trp Ser 610 615 620	1870
GAT GTG ACC ACC AGC CCT ACC CCT GGG GGT GCT GGA GGA GGT CAG CGG Asp Val Thr Thr Ser Pro Thr Pro Gly Gly Ala Gly Gly Gln Arg 625 630 635	1918
GGA CCT GGG ACA CAC ATG TCT GAG CCA CCA CAC AAC AAC ATG CAG GTT Gly Pro Gly Thr His Met Ser Glu Pro Pro His Asn Asn Met Gln Val 640 645 650 655	1966
TAT GCG TGAGAGAGTC CACCTCCAGT GTAGAGACAT AACTGACTTT TGTAAATGCT Tyr Ala	2022
GCTGAGGAAC AAATGAAGGT CATCCGGGAG AGAAATGAAG AAATCTCTGG AGCCAGCTTC	2082
TAGAGGTAGG AAAGAGAAGA TGTCTTATT CAGATAATGC AAGAGAAGCA ATTGTCAGT	2142
TTCACTGGGT ATCTGCAAGG CTTATTGATT ATTCTAATCT AATAAGACAA GTTGTGGAA	2202
ATGCAAGATG AATAACAAGCC TTGGGTCCAT GTTACTCTC TTCTATTGG AGAATAAGAT	2262
GGATGCTTAT TGAAGCCCAG ACATTCTGC AGCTTGGACT GCATTTAAG CCCTGCAGGC	2322
TTCTGCCATA TCCATGAGAA GATTCTACAC TAGCGTCCTG TTGGGAATTAA TGCCCTGGAA	2382
TTCTGCCTGA ATTGACCTAC GCATCTCCTC CTCCCTGGAC ATTCTTTGT CTTCATTTGG	2442
TGCTTTGGT TTTGCACCTC TCCGTGATTG TAGCCCTACC AGCATGTTAT AGGGCAAGAC	2502
CTTTGTGCTT TTGATCATTG TGGCCATGA AAGCAACTT GGTCTCCTT CCCCTCCTGT	2562
CTTCCCGGTA TCCCTTGGAG TCTCACAGG TTTACTTTGG TATGGTTCTC AGCACAAACC	2622
TTTCAAGTAT GTTGTCTT TGGAAAATGG ACATACTGTA TTGTGTCTC CTGCATATA	2682
CATTCCCTGGA GAGAGAAGGG GAGAAGAATA CTTTCTTCA ACAAAATTG GGGGCAGGAG	2742
ATCCCTCAA GAGGCTGCAC CTTAATTCTT CTTGTCTGTG TGCAGGTCTT CATATAAACT	2802

FIG.11E

TTACCAAGAA GAAGGGTGTG AGTTTGTGT TTTCTGTGT ATGGGCCTGG TCAGTGTAAA	2862
GT ₄ TATCCT TGATAGTCTA GTTACTATGA CCCTCCCCAC TTTTTAAAAA CCAGAAAAAG	2922
GT ₄ TGGAATG TTGGAATGAC CAAGAGACAA GTTAACTCGT GCAAGAGCCA GTTACCCACC	2982
CACAGGTCCC CCTACTTCCT GCCAAGCATT CCATTGACTG CCTGTATGGA ACACATTGT	3042
CCCAGATCTG AGCATTCTAG GCCTGTTCA CTCACTCACC CAGCATATGA AACTAGTCTT	3102
AACTGTTGAG CCTTCCTT CATA ₂ CCACA GAAGACACTG TCTCAAATGT TGTACCC ₂ TG	3162
CCATTAGGA CTGAAC ₂ TTC CTTAGCCCAA GGGACCCAGT GACAGTTGTC TTCCGTTGT	3222
CAGATGATCA GTCTCTACTG ATTATCTTGC TGCTAAAGG CCTGCTCACC AATCTTCTT	3282
TCACACCGTG TGGTCCGTGT TACTGGTATA CCCAGTATGT TCTCACTGAA GACATGGACT	3342
TTATATGTTA AAGTGCAGGA ATTGGAAAGT TGGACTTGTT TTCTATGATC CAAAACAGCC	3402
CTATAAGAAG GTTGGAAAAG GAGGAACATAT ATAGCAGCCT TTGCTATTTT CTGCTACCAT	3462
TTCTTTCCCT CTGAAGCGGC CATGACATT ₂ CTTTGGCAA CTAACGTAGA AACTCAACAG	3522

FIG.11F

AACATTTCC TTTCTAGAG TCACCTTTA GATGATAATG GACAACATA GACTTGCTCA 3582
TTGTTCAGAC TGATTGCCCC TCACCTGAAT CCACTCTCTG TATTCATGCT CTTGGCAATT 3642
TCTTGACTT TCTTTAAGG GCAGAAGCAT TTTAGTTAAT TGTAGATAAA GAATAGTTT 3702
CTTCCTCTC TCCTTGGGCC AGTTAATAAT TGGTCCATGG CTACACTGCA ACTTCCGTCC 3762
AGTGCTGTGA TGCCCAGTAC ACCTGCAAAA TAAGTTCTGC CTGGGCATTT TGTAGATATT 3822
AACAGGTGAA TTCCCAGACTC TTTGGTTTG AATGACAGTT CTCATTCTT CTATGGCTGC 3882
AAGTATGCAT CAGTGCTTCC CACTTACCTG ATTTGTCTGT CGGTGGCCCC ATATGGAAAC 3942
CCTGCCTGTC TGTTGGCATA ATAGTTACA AATGGTTTT TCAGTCCTAT CCAAATTAT 4002
TGAACCAACA AAAATAATTA CTTCTGCCCT GAGATAAGCA GATTAAGTTT GTTCATTCTC 4062
TGCTTTATTC TCTCCATGTG GCAACATTCT GTCAAGCCTCT TTCATAGTGT GCAAACATT 4122
TATCATTCTA AATGGTGACT CTCTGCCCTT GGACCCATT ATTATTACA GATGGGGAGA 4182
ACCTATCTGC ATGGACCCCTC ACCATCCTCT GTGCAGCACA CACAGTGCAG GGAGCCAGTG 4242
GCGATGGCGA TGACTTTCTT CCCCTG 4268

FIG. 11G

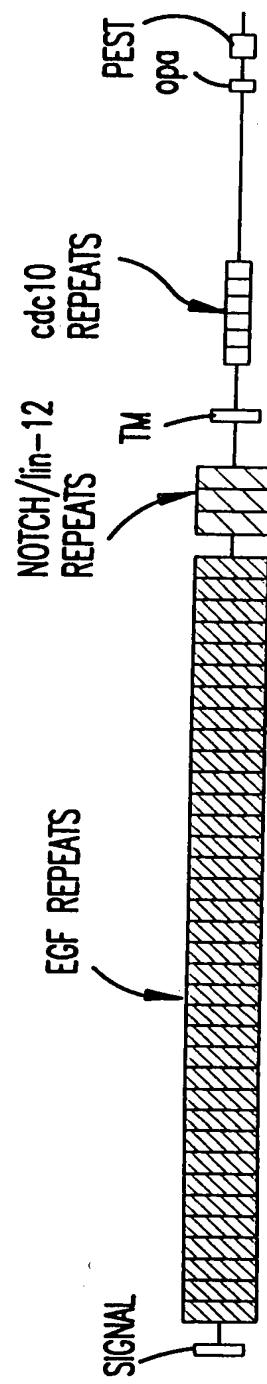


FIG.12A

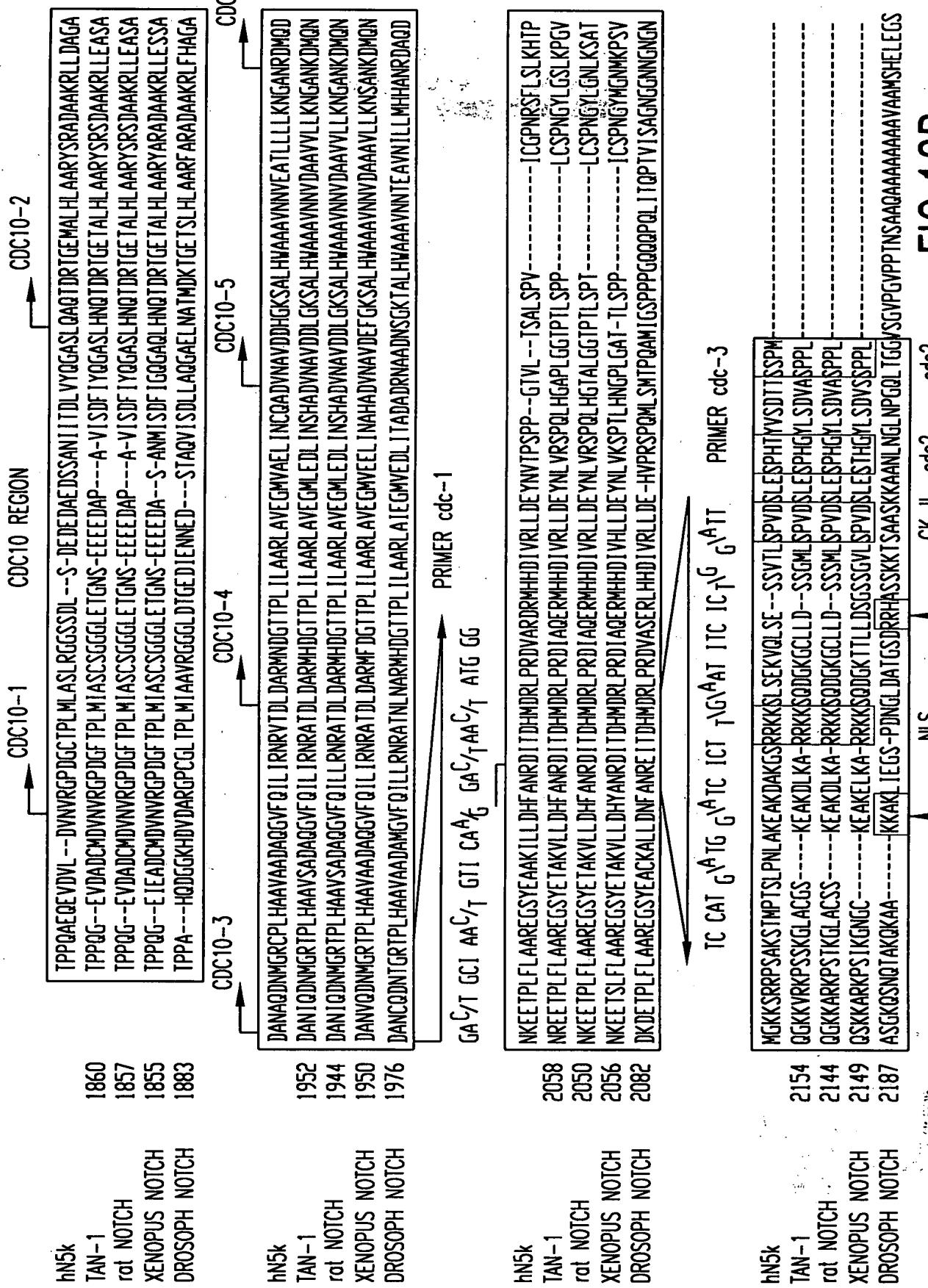


FIG. 12B

42

hnsk			
TAN-1	2218		
rat NOTCH	2209		
XENOPUS NOTCH	2214		
DROSOPH NOTCH	2285	PVGGMGGNLPSPYDTSSMYSNAMAAPLAGNPNTGAKQPRSYEDCIKNAQSMQSLQNGNLDMIKLDNAYASMCSPR	CK-II
hnsk			
TAN-1	2250	A-KPEMAALGGGRGLAFTGPPRLSHPVASSGTSTVLSQGGNLFNFTGGGSTSLLNGQCEWL	PLAHCASTVLPVSQQLLSHHHVSPGS
rat NOTCH	2242	A-KPEMAALAGGSRLAEEPPPRRLSHPVASSAATVLSLNGQCEWL	-GSAGSSLRSLHPVVPADW
XENOPUS NOTCH	2247	T-KQEMAA--GSNRMWAFDAMVPRLTHL-NASSPNTIMS	-MNRMEMVNETQYNEMFGMVLAEG-THPGI
DROSOPH NOTCH	2390	GGLCGMGGLSGAGGNGSHEQGLSPPPYS-NQSPPHSVAQSSLALSPHAYLGSPPSAKSRPSSLPTSPTHI	QQELLNGQGLGMNGQQRNGVGPGLP

hnsk			
TAN-1	2354	AASALSQMMS	APQPOSTCPPAVAGPLPTMVP
rat NOTCH	2344	STNTLSPTI	APQPOSTCPPAVAGPLPTMVP
XENOPUS NOTCH	2343	PATLSSQMMT	APQPOSTCPPAVAGPLPTMVP
DROSOPH NOTCH	2495	VGGQQPQNSPVSLLGII	APQPOSTCPPAVAGPLPTMVP
hnsk			
TAN-1	2448	SSLAVHTLPQ-ESPALPTSLPSSLVPPVTAAQFLTPPSQHSY-SS-PVENTPSQHSY-SS-PMNTPSQHSY-SS-PMNTPSQHSY-SS	APQPOSTCPPAVAGPLPTMVP
rat NOTCH	2423	SSLAVHTLPQ-ESQALPTSLPSSLVPPVTAQFLTPPSQHSY-SS-PMNTPSQHSY-SS-PMNTPSQHSY-SS-PMNTPSQHSY-SS	APQPOSTCPPAVAGPLPTMVP
XENOPUS NOTCH	2416	SSNNIHSYMPQ-DTQIFAAASLPSNLTQSMTAQFLTPPSQHSY-SS-PMNTPSQHSY-SS-PMNTPSQHSY-SS-PMNTPSQHSY-SS	APQPOSTCPPAVAGPLPTMVP
DROSOPH NOTCH	2599	S---IASSMSG-SSSPSTNMLSPSSQHNQQAFYQYLTPSSQHS	APQPOSTCPPAVAGPLPTMVP

hnsk			
TAN-1	2448	SSLAVHTLPQ-ESPALPTSLPSSLVPPVTAQFLTPPSQHSY-SS-PVENTPSQHSY-SS-PMNTPSQHSY-SS-PMNTPSQHSY-SS	SVAFTAMMPQQDQVQATILPAYHFPASVGKHTPPSQHSYASSNAAERTPSHSCHIAGEHPYLTPSPESPDQWSSSSPSHSA-SDIVSDVTSPTP
rat NOTCH	2423	SSLAVHTLPQ-ESQALPTSLPSSLVPPVTAQFLTPPSQHSY-SS-PMNTPSQHSY-SS-PMNTPSQHSY-SS-PMNTPSQHSY-SS	SSLAVHTLPQ-ESQALPTSLPSSLVPPVTAQFLTPPSQHSY-SS-PMNTPSQHSY-SS-PMNTPSQHSY-SS-PMNTPSQHSY-SS
XENOPUS NOTCH	2416	SSNNIHSYMPQ-DTQIFAAASLPSNLTQSMTAQFLTPPSQHSY-SS-PMNTPSQHSY-SS-PMNTPSQHSY-SS-PMNTPSQHSY-SS	SSLAVHTLPQ-ESQALPTSLPSSLVPPVTAQFLTPPSQHSY-SS-PMNTPSQHSY-SS-PMNTPSQHSY-SS-PMNTPSQHSY-SS
DROSOPH NOTCH	2599	S---IASSMSG-SSSPSTNMLSPSSQHNQQAFYQYLTPSSQHS	SSLAVHTLPQ-ESQALPTSLPSSLVPPVTAQFLTPPSQHSY-SS-PMNTPSQHSY-SS-PMNTPSQHSY-SS-PMNTPSQHSY-SS

PEST-CONTAINING REGION

FIG. 12C

Potential signal cleavage site

hum N	MP	ALRPAL LWALLALWLC CA	APA HA	
TAN-1	MP	PL LAPLLCLALL PA	LAA RG	P
Xen N	MD	RIGLAVLLCS LP	VLT QC	L
Dros N	MQSQRSSRRRS RAPNTWICFW INKMHAVASL PASLPLLLT LAFANLPNIV RGTDTALVAA			

hum N	MLGKATCRCA SGFTGEDCQY STSHPCFVSR PCLNGGTCHM LSRDT-YECT CQVGF TGKEC
TAN-1	GVADYACSCA LGFSGPLCLT PLDNAC-LTN PCRNGGTCDL LT-LTEYKCR CPPGWSKGSC
Xen N	NAIDFICHCP VGFTDKVCLT PVDNAC-VNN PCRNGGTCEL LNSVTEYKCR CPPGWTGDSC
Dros N	GRPGISCKCP LGFDESCLCEI AVPNAC-DHV TCLNGGTSQL KT-LEEYTCA CANGYTGERC

hum N	NLPGSYQCQC PQGFTGQYCD SLYVPCAPSP CVNGGTCRQT GDFTFECNCL PGFEGSTCER
TAN-1	NEVGSYRCVC RATHTGPNCE RPYVPCSPSP CQNGGTCRPT GDVTHECACL PGFTGQNCEE
Xen N	NEFGSYRCTC QNRFTGRNCD EPYVPCNPSP CLNGGTCRQT DDTSYDCTCL PGFSGQNCEE
Dros N	NTHGSYQCMC PTGYTGKDCD TKYNPCSPSP CQNAGICRSN G-LSYECKCP KGFEKNCEQ

EGF-like Repeats

QCRDGYEPBV NEGMCVTYHN GTGYCKCPEG FLGEYCQHRD PCE-KNRCQN GGTC—VAQA	83
RCSQPGETCL NGGKCEA-AN GTEACVCGGA FVGPRCQDPN PCL-STPCKN AGTCHVVDRR	80
RCTQTAEMCL NGGRCEMTPG GTGVCLCGNL YFGERCQFPN PCTIKNQCMN FGTCEPVLQG	90
SCTSVG—CQ NGGTCVTQLN GKTYCACDSH YVGDYCEHRN PCN-SMRCQN GGTCAQVTFRN	117
QWTDACLSHP CANGSTCTTV —ANQFSCKC LTGFTGQKCE TDVNEC-DIP GHQCHGGTCL	199
QQADPCASNP CANGQCLPF —EASYIHC PPSFHGPTCR QDVNECGQKP RLCRHGGTCH	196
QQADPCASNP CANGGKCLPF —EIQYICKC PPGFHGATCK QDINEC-S-Q NPCKNGGQCI	195
ETKNLCASSP CRNGATCTAL AGSSSTCSC PPGFTGDTCS YDIEEC-Q-S NPCKYGGICV	233
NIDDCPNHRC QNGGVCVDGV NTYNCRCPPQ WTGQFCTEDV DECLLQPNA- CQNGGTCANR	318
NIDDCPGNNC KNGGACVDGV NTYNCPCPPE WTGQYCTEDV DECQLMPNA- CQNGGTCHNT	315
NIDDCPSNNC RNGGTCVDGV NTYNCQCPPD WTGQYCTEDV DECQLMPNA- CQNGGTCHNT	314
NYDDCLGHLC QNGGTCIDGI SDYTCRCPPN FTGRFCQDDV DECAQRDHPV CQNGATCTNT	352

FIG.13A

hum N NGGYGCVVN GWSGDDCSEN IDDCAFASCT PGSTCIDRVA SFSCMCPEGK AGLLCHLDDA
TAN-1 HGGYNCVNV GWTGEDCSEN IDDCASAACF HGATCHDRVA SFYCECPHGR TGLLCHLND
Xen N YGGYNCVNV GWTGEDCSEN IDDCANAACH SGATCHDRVA SFYCECPHGR TGLLCHLDNA
Dros N HGSYSCICVN GWAGLDCSNN TDDCKQAACF YGATCIDGVG SFYCQCTKGK TGLLCHLDDA

hum N AFHCECLKGY AGPRCEMDIN ECHSDPCQND ATCLDKIGGF TCLCMPGFKG VHCELEINEC
TAN-1 SFECQCLQGY TGPRCEIDVN ECVSNPCQND ATCLDQIGEF QCICMPGYEG VHCEVNTDEC
Xen N SFQCNCQPGY AGPRCEIDVN ECLSNPCQND STCLDQIGEF QCICMPGYEG LYCETNIDEC
Dros N SYRCNCNSQGF TGPRCETNIN ECESHPCQNE GSCLDDPGTF RCVCMPCFTG TQCEIDIDEC

hum N ATGFTGVLC EIIDNCDPDP CHHGQCQDG I DSYTCICNPG YMGAICSDQI DECYSSPCLN
TAN-1 TEGYTGTCE VDIDECDPPD CHYGSCKDG V ATFTCLCRPG YTGHHCETNI NECSSQPCRL
Xen N TEGFTGRHCE QDINECIPDP CHYGTCKDG I ATFTCLCRPG YTGRLCDNI NECLSKPCLN
Dros N PPGYTGTSC E ININDCDSNP CHRGKCIDDV NSFKCLCDPG YTGYICQKQI NECESNPCQF

CISNPCHKGA LCDTNPLNGQ YICTCPQGYK GADCTEDVDE CAMANSNPCE HAGKCVNTDC
CISNPCNEGS NCDTNPVNGK AICTCPGTYT GPACSQDVDE CSLG-ANPCE HAGKCIINTLG
CISNPCNEGS NCDTNPVNGK AICTCPGTYT GPACNNDVDE CSLG-ANPCE HGGRCTNTLG
CTSNPCHADA ICDTSPINGS YACSCATGYK GVDCSEDIDE CDQG-SPCE HNGICVNTPG

438
434
433
470

QSNPCVNNQ CVVKVNRFQC LCPPGFTGPV CQIDIDDCSS TPCLNGAKCI DHPNGYECQC
ASSPCLHNGR CLDKINEFQC ECPTGFTGHL CQYDVDECAS TPCKNGAKCL DGPNTYTCVC
ASNPCLHNGK CIDKINEFRC DCPTGFSGNL CQHDFDECTS TPCKNGAKCL DGPNSYTCQC
QSNPCLNDGT CHDKINGFKC SCALGFTGAR CQINIDDCQS QPCRNRGICH DSIAGYSCEC

558
554
553
590

DGRCIDLVNG YQCNCQPGTS GVNEINFDD CASNPCIHG- ICMDGINRYS CVCSPGFTGQ
RGTQDPDNA YLCFCLKTT GPNCEINLDD CASSPCDSG- TCLDKIDGYE CACEPGYTGS
GGQCTDRENG YICTCPKGTT GVNCETKIDD CASNLCDNG- KCIDKIDGYE CTCEPGYTGK
DGHQDRVGS YYCQCQAGTS GKNCEVNNE CHSNPCNNGA TCIDGINSYK CQCVPGFTGQ

677
673
672
710

FIG.13B

hum N	RCNIDIDECA SNPCRKGATC INGVNGFRCI CPEGPHHPSC YSQVNECLSN PCI-HGNCTG
TAN-1	MCNSNIDECA GNPCHNGGTC EDGINGFTCR CPEGYHDPTC LSEVNECNSN PCV-HGACRD
Xen N	LCNININECD SNPCRNGGTC KDQINGFTCV CPDGYHDHMC LSEVNECNSN PCI-HGACHD
Dros N	HCEKNVDECI SSPCANNNGVC IDQVNGYKCE CPRGFYDAHC LSDVDECASN PCVNEGRCED
hum N	DECASNPLCN QGTCFDDISG YTCHCVLPYT GKNCQTVLAP CSPNPCENAA VCKESPNFES
TAN-1	NECASNPCLN KGTCIDDVAG YKCNCLLPYT GATCEVVLAP CAPSPCRNGG ECRQSEDYES
Xen N	NECSSNPCLN HGTCIDDVAG YKCNCMLPYT GAICEAVLAP CAGSPCKNGG RCKESEDFET
Dros N	DDCVTNPCGN GGTCIDKVNG YKCVCKVPFT GRDCESKMDP CASNRCKNEA KCTPSSNFLD
hum N	CLANPCQNGG SCMDGVNTFS CLCLPGFTGD KCQTDNMECL SEPCKNGGTC SDYVNSYTCK
TAN-1	CRPNPCHNGG SCTDGINTAF CDCLPGFRGT FCEEDINECA SDPCRNGANC TDCVDSYTCT
Xen N	CQPNPCHNGG SCSDGINMFF CNCPAGFRGP KCEEDINECA SNPKNGANC TDCVNSYTCT
Dros N	CASFPCQNGG TCLDGIGDYS CLCVDGFDGK HCETDINECL SQPCQNGATC SQYVNSYTCT

GLSGYKCLCD AGWVGINCEV DKNECLSNPC QNGGTCNDLV NGYRCTCKKG FKGYNCQVNI	796
SLNGYKCDCD PGWSCTNCDI NNNECESNPC VNGGTCKDMT SGIVCTCREG FSGPNCQTNI	792
GVNGYKCDCE AGWGSNCDI NNNECESNPC MNGGTCKDMT GAYICTCKAG FSGPNCQTNI	791
GINEFICHCP PGYTGKRCEL DIDECSSNPC QHGGTCYDKL NAFSCQCMPG YTGQKCEtni	830
YTCLCA-PGW QGQRCTIDID EC-ISKPCMN HGLCHNTQGS YMCECPPGFS GMDCEEDIDD	914
FSCVCPAGA KGQTCEVDIN EC-VLSPCRH GASCNTHGG YRCHCQAGYS GRNCETDIDD	911
FSCECP-PGW QGQTCEIDMN EC-VNRPCRN GATCQNTNGS YKCNCKPGYT GRNCEMDIDD	909
FSCTCK-LGY TGRYCDDEDID ECSVSSPCRN GASCLNVPGS YRCLCTKGYE GRDCAINTDD	949
CQAGFDGVHC ENNINECTES SCFNNGTCVD GINSFTCLCP VGFTCSFCLH EINECSSHPC	1034
CPAGFSGIHC ENNTPDCTES SCFNNGTCVD GINSFTCLCP PGFTGSYCYQH VVNECDSRPC	1031
CQPGFSGIHC ESNTPDCTES SCFNNGTCID GINTFTCQCP PGFTGSYCYQH DINECDSKPC	1029
CPLGFSGINC QTNDEDCTES SCLNGGSCID GINGYNCSCl AGYSGANCQY KLNKCDNSPC	1069

FIG.13C

16

hum N	LNEGTCVDGL GTYRCSCPGL YTGKNCQTLV NLCSRSPCKN KGTCVQKKAE SQCLCPGWA
TAN-1	LLGGTCQDGR GLHRCTCPQG YTGPNCQNLV HWCDSSPCKN GGKCWQTHQ YRCECPGWT
Xen N	LNGGTCQDSY GTYKCTCPQG YTGLNCQNLV RWCDSSPCKN GGKCWQTNMF YRCECKSGWT
Dros N	LNGATCHEQN NEYTCHCPSG FTGKQCSEYV DWCGQSPCEN GATCSQMKHQ FSCKCSAGWT

hum N	SNPCQHGATC SDFIGGYRCE CVPGYQGVNC EYEVDECQNN PCQNGTCID LVNHFKCSCP
TAN-1	PSPCQNGATC TDYLGGYSCK CVAGYHGVCN SEEIDECLSH PCQNGTCID LPNTYKCSCP
Xen N	PNPCQNGATC TDYLGGYSCE CVAGYHGVCN SEEINECLSH PCQNGTCID LINTYKCSCP
Dros N	SQPCQNGGTC RDLIGAYECQ CRQGFQQQNC ELNIDDCAPN PCQNGGTCHD RVMNFSCSCP

hum N	CLSNPCSSEG SLDCIQLTND YLCVCRSAFT GRHCTFVDV CPQMPCLNGG TCAVASNMPD
TAN-1	CLSNPCDARG TQNCVQRVND FHCECRAHT GRRCESVING CKGKPCKNNG TCAVASNTAR
Xen N	CLSNPCDSRG TQNCIQLVND YRCECRQGFT GRRCESVVDG CKGMPCRNNG TCAVASNTER
Dros N	CLSNPCSNAG TLDVQLVNN YHCNCRPGHM GRHCEHKVDF CAQSPCQNGG NCNI—RQS

GAYCDVPNVS CDIAASRRGV LVEHLCQHSG VCINAGNTHY CQCPLGYTGS YCEEQLDECA	1154
GLYCDVPSVS CEVAAQQRQGV DVARLCQHGG LCVDAGNTHH CRCQAGYTG YCEDLVDECS	1151
GYCDVPSVS CEVAAKQQGV DIVHLCRNSG MCVDGTNHF CRCQAGYTG YCEEQVDECS	1149
GKLCDVQTIS CQDAADRKGSLRQLC—NNG TCKDYGNSHV CYCSQGYAGS YCQKEIDEQ	1188
PGTRGLLCEE NIDDCAR— GPHCLN GGQCMDRIGG YSCRCLPGFA GERCEGDINE	1267
RGTQGVHCEI NVDDCNPPVD PVSRSRKCFN NGTCVDQVGG YSCTCPPGFV GERCEGDVNE	1271
RGTQGVHCEI NVDDCTPFYD SFTLEPKCFN NGKCIDRVGG YNCICPPGFV GERCEGDVNE	1269
PGTMGIICEI NKDDCKP— GACHN NGSCIDRVGG FECVCQPGFV GARCEGDINE	1300
GFICRCPPGF SGARCQS— SCGQVKCRKG EQVHTAS— GPRCFCPSP— RDCE—	1376
GFICKCPAGF EGATCENDAR TCGSLRCLNG GTCISGPR— SPTCLCLGPV TGPECQFPAS	1389
GFICKCPPGF DGATCEYDSR TCSNLRCQNG GTCISVLT— SSKVCVSEGY TGATCQYPVI	1387
GHHCICNNGF YGKNCELSGQ DCDSNPCRVG NCVADEGF GYRCECPRT LGEHCEIDTL	1415

FIG.13D

hum N	-GC-ASSPCQ HGGSCHPQRQ PPYYSCQCAP PFSGSCEL	-YTAPP	-S	TPP
TAN-1	SPCLCCNPCY NQGTCCEPTSE SPFYRCLCPA KFNGLLCHIL DYSFGC	-GAGRD	I	PPP
Xen N	SPC-ASHPCY NGGTCQFFAE EPFFQCFCPK NFNGLFCHIL DYEFPG	-GLGKN	I	TPP
Dros N	DEC-SPNPCA QGAACEDLLG D-YECLCPs KWKGKRCDIY DANYPGWNGG SGSGNDRYAA			

hum N	NN-QCDELCN TVECLFDNFE CQGNSKTCK-	-YDKYCADHF	KDNHCNQGCN	SEECGWDGLD
TAN-1	SDGHCDSQCN SAGCLFDGFD CQRAEGQCNP LYDQYCKDHF	SDGHCDQGCN	SAECEWDGLD	
Xen N	NDGKCDSQCN NTGCLYDGFD CQKVEVQCNP LYDQYCKDHF	QDGHCDQGCN	NAECEWDGLD	
Dros N	KNGKCNEECN NAACHYDGHD CERKLKSCDS LFDAYCQKHY GDGFCDYGCN NAECSDWGLD			

hum N	YYGEKSAAMK KQ-R	MTRRSL	PCEQ	E QEVAGSKVFL
TAN-1	YYGREEELRK HPIKRAAEW AAPDALLGQV KASLLPGGSE GGRRRRELDP MDVRGSIVYL			
Xen N	YYGNEEELKK HHIKRSTDYW SDAPSAI	-FSTMKESIL	LGRHRRELDE	MEVRGSIVYL
Dros N	WKDNVRVPEI EDTDFARKNK ILYTQQVHQ			TGIIQIYL

LNR (Notch/Lin-12 Repeats)

—A—	TCL SQYCADKARD GVCDEACNSH ACQWDGGDCS LT MENPWANC SSPLPCWDYI	1476
LIEE—	ACE LPECQEDAGN KVCSLQCNH ACGWDGGDCS LN FNDPWKNC TQSLQCWKYF	1501
DND—	ICE NEQCSELADN KVCNANCNNH ACGWDGGDCS LN FNDPWKNC TQSLQCWKYF	1498
DLEQQRAMCD	KRGCTEKQGN GICDSDCNTY ACNFDGNDCS LGI-NPWANC TAN-EXWNKF	1531

CAADQOPEN-L	AEGTLVIVVL MPPEQLLQDA R-SFLRALGT LLHTNLRIKR DSQGELMVYP	1591
CAEHVPER-L	AAAGTL-VVVV LMPPEQLRNS SFHFLRELSR VLHTNVVFKR DAHGQQMIFP	1619
C-ANMPEN-L	AEGTLVLVVL MPPEQLRNS V-NFLRELSR VLHTNVVFKK DSKGEYKIYP	1615
CENIKTQSPVL	AEGAMSVVML MNVEAFREIQ A-QFLRNMSH MLRTTVRLKK DALGHDIIIN	1650

EIDNRQCVQD	SDHCFKNTDA AAALLASHAI QG-TLSYP LSVVSESLT PERT-Q-LLY	1680
EIDNRQCVQA	SSQCFQSATD VAAFLGALAS LGSL-NIPYK IEAVQSETVE PPPPAQ-LHF	1737
EIDNRQCYKS	SSQCFNSATD VAAFLGALAS LGSLDTLSYK IEAVKSENME TPKPST-LYP	1730
EIDNRKCTEC	FTHAVEAAEF LAATAAKHQL RNDFQ-IHSV RGIKNPGDED NGEPPANVKY	1745

FIG.13E

hum N	LLAVAVVIIL	FIILLGVIMA	KRKRK—HGS	LWLPEGFTLR	RDASNHKRRE	PVGQDAVGLK
TAN-1	MYVAAAFLV	LFFVGGCVL	SRKRRRQHQQ	LWFPEGFKV—	SEASKKKRRE	ELGEDSVGLK
Xen N	MLSMLVIPLL	IIFVFMIV	NKKRRRREHDS	FGSPTALFQK	NPA—KRNGET	PW—EDSVGLK
Dros N	VITGIIILVII	ALAFFGMVL	—STQRKRAHGV	TWFPEGFRAP	AAVMSRRRRD	PHGQEMRNLN

CDC-10/Ankyrin Repeats

hum N	PIDRRPWTQQ HLEAADIRRT PSLALTPPQA EQEVDVLDVN VRGPDGCTPL MLASLRGGSS
TAN-1	QTDHRQWTQQ HLDAADL-RM SAMAPTPPQC EVDADCMDVN VRGPDGFTPL MIASCSGGGL
Xen N	KTDPRQWTRQ HLDAADL-RI SSMAPTPPQC EIEADCMDVN VRGPDGFTPL MIASCSGGGL
Dros N	EADQQRWWSQA HLDVVVDV-R- AIM—TPP-A HQDGGKHDVD ARGPCGLTPL MIAAVRGGGL

hum N	ANQDQDNMGRC	PLHAAVAADA	QGVFQILIRN	RVTLDLADMN	DGTTPLILAA	RLAVEGMVAE
TAN-1	ANIQDQDNMGRT	PLHAAVSADA	QGVFQILIRN	RATLDLADMH	DGTTPLILAA	RLAVEGMLED
Xen N	ANVQDQDNMGRT	PLHAAVAADA	QGVFQILIRN	RATLDLADMF	DGTTPLILAA	RLAVEGMVEE
Dros N	ANCQDQNTGRT	PLHAAVAADA	MGVFQILLRN	RATNLNARMH	DGTTPLILAA	RLAIEGMVED

NLSVQVSEAN	LIGTGTSEHW	VDDE	—	—	—	G	PQPKKVKAED	EALLSE-EDD	1782
PLK-NASDGA	LMDDNQNE-W	GDED	—	—	—	—	LETKKFRFEE	PVVLPD-LDD	1837
PIK-NMTDGS	FMDDNQNE-W	GDEET	—	—	—	—	LENKFRFEE	QVILPELVDD	1831
KQVAMQSQGV	GQPGAH-W	SDDE	SDMPLP	KRQRSDPVSG	VGLGNNGGYA	SDHTMSEYE	—	—	1861

DLSDEDEDAE	DSSANI	ITDL	VYQGASLQAQ	TDRTGEMALH	LAARYSRADA	AKRLLDAGAD	1902
ETGNSEEE-E	DAPA-VISDF		IYQGASLHNQ	TDRTGETALH	LAARYSRSDA	AKRLLLEASAD	1954
ETGNSEEE-E	DASANMISDF		IGQGAQLHNQ	TDRTGETALH	LAARYARADA	AKRLLLESSAD	1949
DTGEDIENNE	DSTAQV	ISDL	LAQGAELNAT	MDKTGETSLH	LAARFARADA	AKRLLDAGAD	1976

LIN	QCADVNA	VDDHGKSALH	WAAAVNNVEA	TLLLKNGAN	RDMQDNKEET	PLFLAAREGS	2022
LIN	SHADVNA	VDDLGSKALH	WAAAVNNVDA	AVVLLKNGAN	KDMQNNREET	PLFLAAREGS	2074
LIN	AHADVNA	VDEFGKSALH	WAAAVNNVDA	AAVLLKNSAN	KDMQNNKEET	SLFLAAREGS	2069
LIT	ADADADINA	ADNSGKTAIH	WAAAVNNTEA	VNIIIMHHAN	RDAODDKDFT	PLFLAAREGS	2096

FIG. 13F

hum N YEAAKILLDH FANRDITDHM DRLPRDVARD RMHHDIVRLL DEYNVTPSPP —GTVL—TS
 TAN-1 YETAKVLLDH FANRDITDHM DRLPRDIAQE RMHHDIVRLL DEYNLVRSPQ LHGAPLGGTP
 Xen N YETAKVLLDH YANRDITDHM DRLPRDIAQE RMHHDIVHLL DEYNLVKSPT LHNGPLGAT-
 Dros N YEACKALLDN FANREITDHM DRLPRDVASE RLHHDIVRLL DE-HVPRSPQ MLSMTPQAMI

	NLS	CK II	cdc2	cdc2
hum N	GSRRKKSLSE KVQLSE—SS VTLSPVDSLE	SPHTYVSDTT	SSPM	
TAN-1	A—RRKKSQDG KGCLLD—SS GMLSPVDSLE	SPHGYLSDVA	SPPL	
Xen N	A—RRKKSQDG KTTLLDSGSS GVLSPVDSLE	STHGYLSDVS	SPPL	
Dros N	GS—PDNGLDA TGSLRRKASS KKTSAASKKA ANLNLNGNPGQ LTGGVSGVPG VPPTNSAAQA BNTS			

hum N	ITSPGILQAS PNPML—ATA APPAPVHAQH	
TAN-1	LPSPF	QQS PSVPLNHLPG MPDTHLGIGH
Xen N	MTSPF	QQS PSMPLNHLTS MPESQLGMNH
Dros N	YEDCIKNAQS MQSLQGNGLD MIKLDNYAYS MGSPF	QQE LLNGQGLGMN GNGQRNGVGP
CK II	cdc2	

ALSPV	ICGP NRSFLSLKHT PMGKKSRRPS AKSTMPTSLP NLAKEAKDAK	2127
TLSPP	LCSP NGYLGSLKPG VQGKKVRKPS SKGLACCS— KEAKDLK	2178
TLSPP	ICSP NGYMGNMKPS VQSKKARKPS IKGNCC— KEAKELK	2170
GSPPPGQQQP	QLITQPTVIS AGNGGNNGNG NASGKQSNQT AKQKAA— KKAKLIE	2208
		2169
		2219
		2213
AAAAAAVAA	MSHELEGSPV GVGGMGNLPS PYDTSSMYSN AMAAPLANCN PNTGAKQPPS	2327
ALSF SNLHEM Q	—PLAHGASTV LPSVSQLLSH HHIVSPGS—	2235
LNVA—KPEM AALGGGGRLA FETGPPRLSH LPVASGTSTV LGSSSGALN FTVGGSTSLN		2306
INMAT—KQEM AA—CSNRMA FDAMVPRLTH L—NASSPNTI MS—NGSMH FTVGGAPTMN		2294
GVLPGGLCGM GGLSGAGNGN SHEQGLSPPY SNQSPPHSVQ SSLALSPHAY LCSPSPAKSR		2445

FIG.13G

hum N GSAGSLSRLH PVPVPADW— MNRMEVNETQ YNEMFGMVL A PAEG—THPGI APQSRPPECK
 TAN-1 GQCEWLSRLQ SCMVNPQYNP LRCVAPGPL STQAPSLQHG —MVGPLHSSL AASALSQMS
 Xen N SQCDWLARLQ NGMVQNQYDP IRNGIQQGN— AQQAQALQHG LMTS—LHNGL PATTLSQMMT
 Dros N PSLPTSPTHI QAMRHATQQK QFGGSNLNSL LGGANGGGVV GGGGGGGGV GQGPQNSPVS

hum N APQPQSTCPP AVAGPLPTMY QIP— EM ARL—PSVAFP TAMMPQQDGQ VAQTILPAYH
 TAN-1 PPQPHLGVSS AASGHLGRSF LSGEPSQADV QPLGPSSLAV HTILPQ—ESP ALPTSLPSSL
 Xen N MQQQHHN—SS TTSTHINSPF CSSDISQTDL QQM—SSNNI HSVMPQ—DTQ IFAASLPSNL
 Dros N QQQLGGLEFG SAGLDLNG—F CGSPDSFHSG QMNPPS—I QSSMSG—SSP STNMLSPSSQ

hum N SDWSDVTTSP TPGGAGGGQR GPGTHMSEPPHNN MQVYA
 TAN-1 SDWSEGVSSP PT— SMQ SQIARIPEAFK
 Xen N SDWSEGISSP PT— SMQ PQRTHIPEAFK
 Dros N SDWSEGVQSP AANNLYISGG HQANKGSEAIYI

—HITTPRE PLPP—IV—TF QLIPKGSIAQ PAG— 2320
 —YQGLPSTRL ATQPHLVQTQ QVQPQNLQMQ QQLQQLQPAQ QQQLQQLPPP 2414
 —YQAMPNTRL ANQPHLMQAQ QMQQQQN— LQLHQ 2384
 LGIISPTGSD MGIMLAPPQS SKNSAIMQTI SPQQQQQQQQ QQQQQHQQQQ QQQQQQQQQ 2565

PEST -containing Region

PFPASVGKYP	TPPSQHSYAS SNAERTPSH SGHLQGEHPY LTPSPESPDQ WSSSSPHSA—	2433
VPPVTAQFL	TPPSQHSY—S S—PVENTPSH QLQVP—EGPF LTPSPESPDQ WSSSSPHSNV	2530
TQSMTTAQFL	TPPSQHSY—S S—PMNDNTPSH QLQVP—DHPF LTPSPESPDQ WSSSSPHSNM	2497
HNQQAFYQYL	TPSSQHS— CGHTPQH LVQTL—D—SY PTPSPESPGH WSSSSPRSN—	2671

2471
 2556
 2523
 2703

FIG.13H

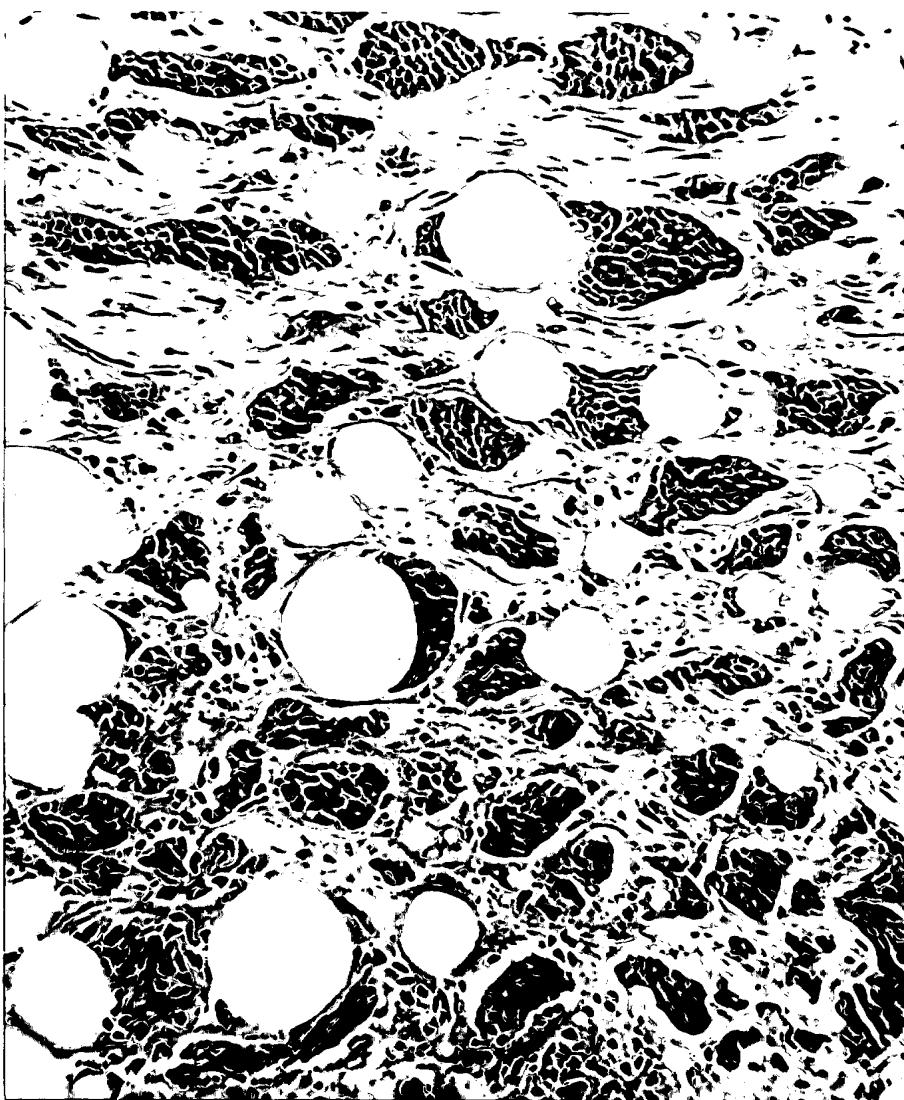


FIG.14

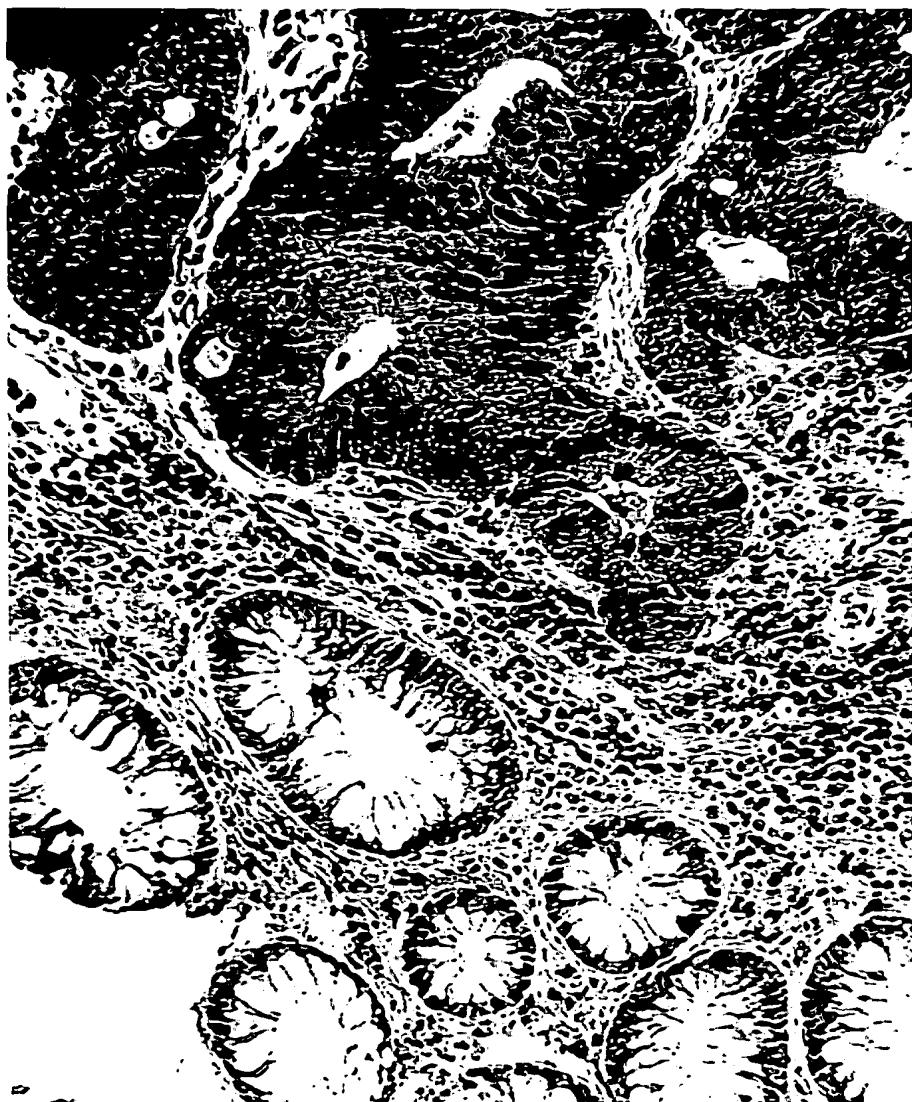


FIG.15A



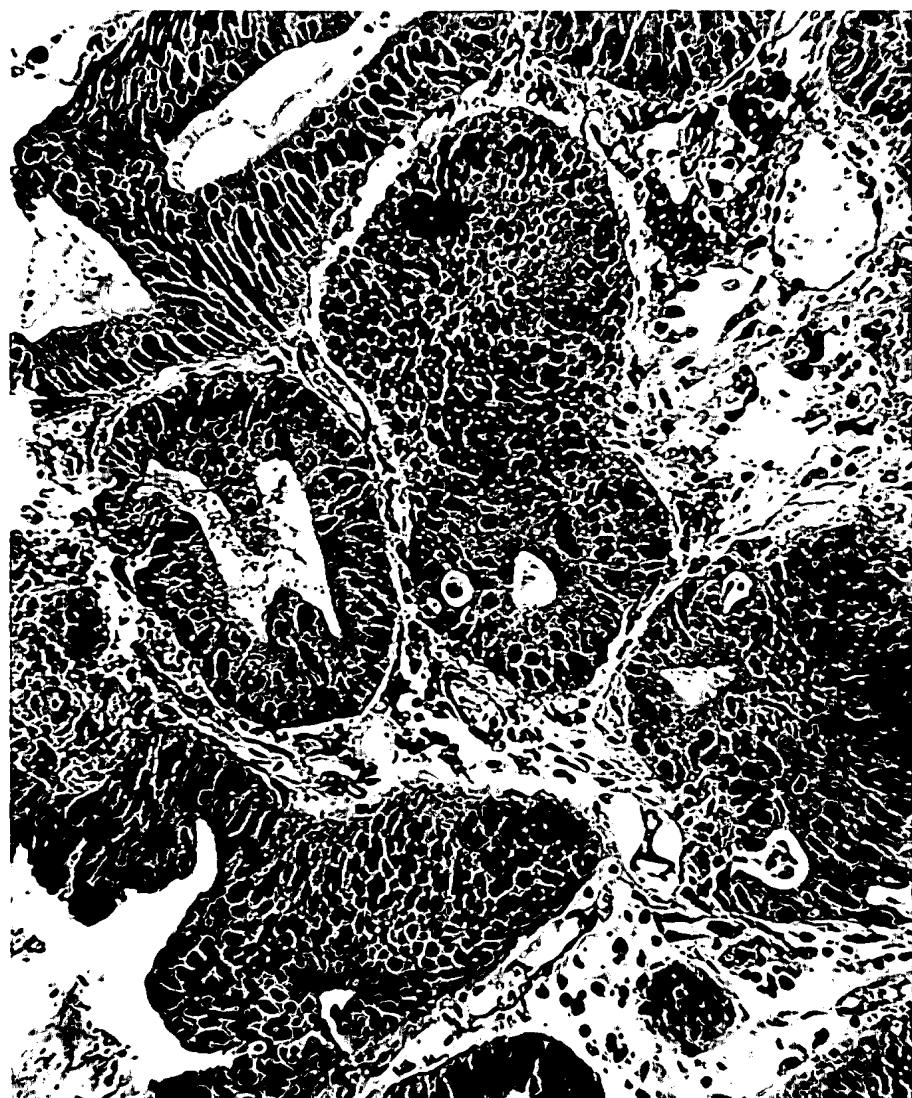


FIG. 15B





FIG.16A





FIG. 16B



10 20 30 40 50 60 70 80 90
 CCAATTCCCC CGGCCCTGCC CCCCGCTCTG CTGTCGGGCCCTCTGGGTGTGC TGCGGGGCCCGCCGGCATGC ATTGCACTGT
 P A L R P A L L W A L L A L W L C C A A P A H A L Q C >
 100 110 120 130 140 150 160 170 180
 CCACATGGCT ATCAACCTG TGTAAATCAA CGAATGTGTG TTACCTACCA CAATGCCACA CGATACTGCA AATGTCCAGA AGGCTTCTTG
 R D G Y E P C V N E G M C V T Y H N G T G Y C K C P E G F L >
 190 200 210 220 230 240 250 260 270
 GGGGAATATT GTCAACATCG AGACCCCTGT GAGAAGAACCGCTGCCAGAA TGCTGGACT TGTGTGGCCCGGCCATGCTGGGAAAGCC
 G E Y C Q H R D P C E K N R C Q N G G T C V A Q A M L C K A >
 280 290 300 310 320 330 340 350 360
 ACCTCCCCAT GTGCCCTCAGG GTTACAGGA GAGGACTGCC ACTACTCAAC ATCTCATCCA TGCTTGTGTG CTGGACCCCTG CCTGAATGCC
 T C R C A S G F T G E D C Q Y S T S H P C F V S R P C L N G >
 370 380 390 400 410 420 430 440 450
 CCCACATGCC ATATGCTCAG CGGGGATACC TATGACTGCA CCTGTCAACT CGGGTTTACA CGTAAGGACT GCAATGGAC CGATGCCCTCC
 G T C H M L S R D T Y E C T C Q V G F T G K E C Q W T D A C >
 460 470 480 490 500 510 520 530 540
 CTCTCTCATC CCTGTGCAAA TGGAACTACC TGTACCACTG TGGCCAACCA GTTCTCTGC AAATGCCCTCA CAGGCTTCAC AGGGCAGAAA
 L S H P C A N G S T C T T V A N Q F S C K C L T G F T G Q K >
 550 560 570 580 590 600 610 620 630
 TGTGAGACTG ATGCAATGA GTGTGACATT CCAGGACACT GCCACCATGG TGCCACCTGC CTCAACCTGC CTGTTCTCA CCACTGCCAG
 C E T D V N E C D I P G H C Q H G C T C L N L P G S Y Q C Q >
 640 650 660 670 680 690 700 710 720
 TGGCCCTCAGG GTTCAACAGG CCAGTACTGT GACACCCCTGT ATGTCCTGC TCCACCCCTCA CCTTGTCTCA ATGGAGGCAC CTGTCGGCAG
 C P Q G F T G Q Y C D S L Y V P C A P S P C V N G G T C R Q >
 730 740 750 760 770 780 790 800 810
 ACTCGTGACT TCACTTTGCA GTGCAACTGCC CTTCCACCTT TTGAAGGGAG CACCTGTGAG AGGAATATTG ATGACTGCCCTAACCACAGG
 T G D F T F E C N C L P G F E G S T C E R N I D D C P N H R >

FIG.17A

820	830	840	850	860	870	880	890	900
TGTCAAAATC GAGGGCTTC TGTGGATGGC GTCAACACTT ACAACTGCCG CTGCCCCA CAATGGACAG GACAGTTCTG CACAGAGGAT								
C Q N	G G V	C V D G	V N T	Y N C R	C P P	Q W T	G Q F C	T E D
910	920	930	940	950	960	970	980	990
GTGGATCAAT GCCTGCTGCA CCCAATGCC TCTCAAATG GGGCACCTG TCCCAACCCC AATGAGGCT ATGGCTGTCT ATCTGTCAC								
V D E	C L L Q	P N A	C Q N	G G T C	A N R	N G G	Y G C V	C V N
1000	1010	1020	1030	1040	1050	1060	1070	1080
GGCTCCACTC GAGATGACTG CAGTGAGAAC ATTGATGATT GTGCCCTCGC CTCCGTACT CCAGGCTCCA CCTCCATCGA CCCGTGCCCC								
G W S	G D D C	S E N	I D D	C A F A	S C T	P G S	T C I D	R V A
1090	1100	1110	1120	1130	1140	1150	1160	1170
TCCTTCTCTT GCATGTGCC AGAGGGAAAG CCACGTCTCC TGTGTCACTT GGATGATGCA TGCATCAGCA ATCCTGCCA CAAGGGGCCA								
S F S	C M C P	E G K	A G L	L C H L	D D A	C I S	N P C H	K G A
1180	1190	1200	1210	1220	1230	1240	1250	1260
CTGTGTGACA CCAACCCCT AAATGGCAA TATATTGCA CCTGCCACA AGCTACAAA GGGCTGACT GCACACAAGA TGTGGATGAA								
L C D	T N P L	N G Q	Y I C	T C P Q	G Y K	G A D	C T E D	V D E
1270	1280	1290	1300	1310	1320	1330	1340	1350
TGTCCCATGG CCAATACCAA TCCTTCTGAG CATGCCAGAA AATGTTGAA CACGGATGCC GCCTTCCACT GTGAGTGTCT GAAGGTTAT								
C A M	A N S N	P C E	H A G	K C V N	T D G	A F H	C E C L	K G Y
1360	1370	1380	1390	1400	1410	1420	1430	1440
GCACGGACCTC GTTGTGAGAT GGACATCAAT GACTGCCATT CAGACCCCTG CCACAAATGAT GCTACCTGTC TCGATAAGAT TGGAGGCTTC								
A G P	R C E M	D I N E	C H S D	P C Q N D	A T C L	D K I	G G F	
1450	1460	1470	1480	1490	1500	1510	1520	1530
ACATGTTGTCGATGCTG CCAATGCCAGG TTTCAACGCT GTCCATTGTC ATTGAAAT AAATGAATGT CAGAGCAACC CTTGTGCAA CAATGCCAC								
T C L	C M P G	F K G	V H C	E L E I	N E C	Q S N	P C V N	N G Q
1540	1550	1560	1570	1580	1590	1600	1610	1620
TGTGTGGATA AACTCAATCC TTCCACTGCC CTGTGTCTC CTGCTTCAC TGGCCAGTT TGCCAGATTG ATATTGATGA CTCTTCCACT								
C V D	K V N R	F Q C	L C P	P G F T	G P V	C Q I	D I D D	C S S

FIG.17B

1630	1640	1650	1660	1670	1680	1690	1700	1710
ACTCCGTGTC	TGAATGGGGC	AAACTCTATC	GATCACCCGA	ATGCCATAGA	ATGCCACTGT	GCCACACGTT	TCACTGCTCT	CTTGTCTGAC
T P C	L N G A	K C I	D H P	N G Y E	C Q C	A T G	F T G V	L C E >
1720	1730	1740	1750	1760	1770	1780	1790	1800
GAGAACATTC	ACAACCTGTGA	CCCCGATCCT	TGCCACCATG	GTCACTGTCA	GGATGGTATT	CATTCCCTACA	CCTCCATCTG	CAATCCCCG
E N I	D N C D	P D P	C H H	G Q C Q	D G I	D S Y	T C I C	N P G >
1810	1820	1830	1840	1850	1860	1870	1880	1890
TACATGGGGC	CCATCTGCAG	TGACCCAGATT	GATCAATCTT	ACAGCACCCCC	TTGCCCTAAC	GATGCCCTGCT	CCATTGACCT	GCTCAATGCC
Y M G	A I C S	D Q I	D E C	Y S S P	C L N	D G R	C I D L	V N G >
1900	1910	1920	1930	1940	1950	1960	1970	1980
TACCACTGCA	ACTGCCAGCC	AGCCACGTCA	GGGGTTAATT	GTCAAATTAA	TTTGATGAC	TGTGCAAGTA	ACCCCTGTAT	CCATGGAATC
Y Q C	N C Q P	G T S	G V N	C E I N	F D D	C A S	N P C I	H G I >
1990	2000	2010	2020	2030	2040	2050	2060	2070
TGTATGGATC	GCATTAATCC	CTACAGTTGT	GTCTGCTCAC	CAGGATTACAC	ACGGCAGAGA	TCTAACATTC	ACATTGATCA	GTGTGCCCTCC
C M D	G I N R	Y S C	V C S	P G F T	G Q R	C N I	D I D E	C A S >
2080	2090	2100	2110	2120	2130	2140	2150	2160
AATCCCTGTC	GCAAGGGTGC	AACATGTATC	AACCGTGTGA	ATGGTTCCG	CTCTATATGC	CCCCAGGGAC	CCCATCACCC	CAGCTGCTAC
N P C	R K G A	T C I	N G V	N G F R	C I C	P E G	P H H P	S C Y >
2170	2180	2190	2200	2210	2220	2230	2240	2250
TCACAGGTGA	ACCAATGCC	GACCAATCCC	TCCATCCATC	AAACTGTAC	TGGAGGTCTC	ACTGGATATA	ACTGTCTCTG	TGATGCAAGCC
S Q V	N E C L	S N P	C I H	G N C T	G G L	S G Y	K C L C	D A G >
2260	2270	2280	2290	2300	2310	2320	2330	2340
TGGCTTGGCA	TCAACTGTGA	ACTGGACAAA	AATGAATGCC	TTTCCAATCC	ATGCCAGAAT	GGAGGAACCT	GTGACAATCT	GGTGAATGGA
W V G	I N C E	V D K	N E C L	S N P	C Q N	G G T	C D N L	V N G >
2350	2360	2370	2380	2390	2400	2410	2420	2430
TACACGTGTA	TTGCAAGAA	GGCCTTAAA	GGCTATACT	GCCACGTGAA	TATTGATGAA	TGTGCCCTCAA	ATCCATGCCCT	GAACCAAGGA
Y R C	T C K F	G F K	G Y N	C Q V N	I D E	C A S	N P C L	N Q G >

FIG.17C

2440	2450	2460	2470	2480	2490	2500	2510	2520
ACCTGCCTTC	ATGACATAAC	TGGCTACACT	TGCCACTGTC	TGCTGCCATA	CACAGGCAAG	AATTGTCAGA	CAGTATTGGC	TCCCTGTTCC
T C F	D D I S	G Y T	C H C	V L P Y	T G K	N C Q	T V L A	P C S >
2530	2540	2550	2560	2570	2580	2590	2600	2610
CCAAACCCCT	GTGAGAATGC	TGCTGTTCC	AAAGACTCAC	CAAATTTCA	GAATTAACT	TGCTTGTC	CTCCCTGGCTC	GCAAGGTCAC
P N P	C E N A	A V C	K E S	P N F E	S Y T	C L C	A P G W	Q G Q >
2620	2630	2640	2650	2660	2670	2680	2690	2700
CGCTGTACCA	TGACATTGA	CGACTGTATC	TCCAAGCCCT	GCATGAACCA	TGGTCTCTGC	CATAACACCC	ACGGCAGCTA	CATCTGTGAA
R C T	I D I D	E C I	S K P	C M N H	G L C	H N T	Q G S Y	M C E >
2710	2720	2730	2740	2750	2760	2770	2780	2790
TCTCCACCAAC	GCTTCAGTGG	TATGGACTGT	GAGGAGGACA	TTGATGACTC	CCTTGCCAAAT	CCTTGCCAGA	ATGGAGGTT	CTGTATGGAT
C P P	G F S G	M D C	E E D	I D D C	L A N	P C Q	N G G S	C M D >
2800	2810	2820	2830	2840	2850	2860	2870	2880
CGACTGAATA	CTTTCTCCCTG	CCTCTGCCCT	CCCCGTTCA	CTGGGGATAA	GTGCCAGACA	CACATGAATC	AGTCTCTGAC	TGAACCCCTGT
G V N	T F S C	L C L	P G F	T G D K	C Q T	D M N	E C L S	E P C >
2890	2900	2910	2920	2930	2940	2950	2960	2970
AACAAATGGAG	GGACCTGCTC	TGACTACGTC	AACAGTTACA	CTTGCAACTC	CCAGGCAGGA	TTTGTGAC	TCCATTGTGA	GAACAAACATC
K N G	G T C S	D Y V	N S Y	T C K C	Q A G	F D G	V H C E	N N I >
2980	2990	3000	3010	3020	3030	3040	3050	3060
AATGACTGCA	CTGAGAGCTC	CTGTTCAAT	GCTGGCACAT	GTCTTGATCC	GATTAACCTCC	TTCTCTTCCT	TCTGCCCTGT	GGCTTCACT
N E C	T E S S	C F N	G G T	C V D G	I N S	F S C	L C P V	G F T >
3070	3080	3090	3100	3110	3120	3130	3140	3150
GGATCCTTCT	GCCTCCATGA	GATCAATGAA	TCCAGCTCTC	ATCCATGCCT	GAATGACCGGA	ACGTGTGTC	ATGCCCTGGC	TACCTACCCG
G S F	C L H E	I N E	C S S	H P C L	N E G	T C V	D G L G	T Y R >
3160	3170	3180	3190	3200	3210	3220	3230	3240
TCCAGGCTGCC	CCCTGGGCTA	CACTGGGAAA	AACTGTCAGA	CCCTGGTGA	TCTCTCCACT	CCGTCTCCAT	GTAAAAACAA	ACGTACTTGT
C S C	P L G Y	T G K	N C Q	T L V N	L C S	R S P	C K N K	G T C >

FIG.17D

3250	3260	3270	3280	3290	3300	3310	3320	3330
CTTCAGAAAA AAGCACACTC CCAGTCCTA TGTCCATCTG GATGGGCTGG TGCCTATTGT GACGTGCCA ATGTCTCTG TGACATACCA V Q K K A E S Q C L C P S G W A G A Y C D V P N V S C D I A>								
3340	3350	3360	3370	3380	3390	3400	3410	3420
GCCTCCAGGA GAGGTGTCT TGTCAACAC TTGCCCCACC ACTCAGGTGT CTGCATCAAT CCTGGCAACA CCCATTACTG TCACTGCC A S R R G V L V E H L C Q H S G V C I N A G N T H Y C Q C P>								
3430	3440	3450	3460	3470	3480	3490	3500	3510
CTGGGCTATA CTGGGAGCTA CTGTGAGGAG CAACTCCATG ACTGTCCGTC CAACCCCTGC CAGCACGGG CAACATGCCAG TGACTTCATT L G Y T G S Y C E E Q L D E C A S N P C Q H G A T C S D F I>								
3520	3530	3540	3550	3560	3570	3580	3590	3600
GGTGGATACA GATGCCACTG TGTCCCAGGC TATCACCGTG TCAACTGTGA CTATGAAGTC GATGAGTCCC AGAACAGCC CTGGCAGAA G G Y R C E C V P G Y Q G V N C E Y E V D E C Q N Q P C Q N>								
3610	3620	3630	3640	3650	3660	3670	3680	3690
GGAGGCACCT GTATTGACCT TGTGAACCAT TTCAAGTGCT CTTGGCCACC AGGCACTCCG GGCCTACTCT GTGAACAGAA CATTGATGAC G C T C I D L V N H F K C S C P P G T R G L L C E E N I D D>								
3700	3710	3720	3730	3740	3750	3760	3770	3780
TGTGCCCGGG GTCCCCATTC CCTTAATGCT CGTCAGTGCA TGGATAGGAT TGGAGGCTAC ACTTGCTCT GCTTGCCTGG CTTTGCCTGG C A R C P H C L N G G Q C M D R I G G Y S C R C L P G F A C>								
3790	3800	3810	3820	3830	3840	3850	3860	3870
GACCGTTGTC AGGGAGACAT CAACGACTCC CTCTCCAACC CCTGCCAGCTC TGAGGGCACC CTGGACTCTA TACACCTCAC CAATGACTAC E R C E G D I N E C L S N P C S S E G S L D C I Q L T N D Y>								
3880	3890	3900	3910	3920	3930	3940	3950	3960
CTCTGTGTTT CCCCTACTGCC CTTTACTGCC CGGCACTGTC AAACCTTCGT CGATGTCTGT CCCCCAGATGC CCTGCCTGAA TGGAGGGACT L C V C R S A F T G R H C E T F V D V C P Q M P C L N G G T>								
3970	3980	3990	4000	4010	4020	4030	4040	4050
TGTGCTGTGG CCAGTAACAT GCCTCATGCT TTCAATTGCC GTTGTCCCCC GGGATTTCC GGGGCAAGGT CCCAGAGGAG CTGTGGACAA C A V A S N M P D G F I C R C P P G F S G A R C Q S S C G Q>								

FIG.17E

4060	4070	4080	4090	4100	4110	4120	4130	4140
GTCAAATCTA GCAACGGGA GCACTGTGTC CACACGGCT CTGGACCCG CTGCTTCGC CCCAGTCCC GGGACTCCG A GTCAGGCTG V K C R K G E Q C V H T A S G P R C F C P S P R D C E S G C >								
4150	4160	4170	4180	4190	4200	4210	4220	4230
GCCACTAGCC CCTGCCAGCA CGGGGGCAGC TGCCACCCCTC AGCCCCAGCC TCCTTATTAC TCCTGCCAGT GTCCCCACC ATTCTCGG A S S P C Q H G G S C H P Q R Q P P Y Y S C Q C A P P F S G >								
4240	4250	4260	4270	4280	4290	4300	4310	4320
ACCGGCTGTC AACTCTACAC GGCAACCCCCC ACCACCCCTC CTGCCACCTG TCTGAGCCAG TATTGTCCCG ACAAGCTCG CGATGGCCTC S R C E L Y T A P P S T P P A T C L S Q Y C A D K A R D G V >								
4330	4340	4350	4360	4370	4380	4390	4400	4410
TGTGATGAGG CCTGCAACAG CCATCCCTGC CACTGGATG GGGGTGACTG TTCTCTCACC ATGGAGAAC CCTGGCCAA CTGCTCC C D E A C N S H A C Q W D G G D C S L T M E N P W A N C S S >								
4420	4430	4440	4450	4460	4470	4480	4490	4500
CCACTTCCCT CCTGGGATTA TATCAACAAAC CAGTGTGATG AGCTGTCAA CACGGTCCAG TGCCTGTTG ACAACTTGA ATGCCAGGG P L P C W D Y I N N Q C D E L C N T V E C L F D N F E C Q G >								
4510	4520	4530	4540	4550	4560	4570	4580	4590
AACAGCAAGA CATGCAACTA TGACAAATAC TGTGGAGACC ACTTCAAAGA CAACCACTGT AACCAGGGT GCAACAGTGA GGAGTGTG N S K T C K Y D K Y C A D H F K D N H C N Q G C N S E E C G >								
4600	4610	4620	4630	4640	4650	4660	4670	4680
TGGGATGGCC TGGACTGTGC TGCTGACCAA CCTGAGAACCC TGGCAGAAGG TACCTGGTT ATTGTGGTAT TGATGCCACC TGAACAACTC W D G L D C A A D Q P E N L A E G T L V I V V L M P P E Q L >								
4690	4700	4710	4720	4730	4740	4750	4760	4770
CTCCACCATG CTGGCAGCTT CTGGGGCA CTGGTACCC TGCTCCACAC CAACCTGCC ATTAAAGGGG ACTCCACGG GCAACTCATC L Q D A R S F L R A L G T L L H T N L R I K R D S Q G E L M >								
4780	4790	4800	4810	4820	4830	4840	4850	4860
GTGTACCCCT ATTATGCTGA GAACTCACCT GCTATGAAGA AACAGAGGAT GACACGGAGA TCCCTTCCTG GTGAACAAAGA ACACGGCTG V Y P Y Y G E K S A A M K K Q R M T R R S L P G E Q E Q E V >								

FIG.17F

4870	4880	4890	4900	4910	4920	4930	4940	4950																						
GCTGGCTCTA AACTCTTCT GGAAATTGAC AACCCCCAGT CTGTTCAAGA CTCACACCAC TGCTTCAAGA ACACGGATCC ACCACGACCT																														
A	G	S	K	V	F	L	E	I	D	N	R	Q	C	V	Q	D	S	D	H	C	F	K	N	T	D	A	A	A	A	>
4960	4970	4980	4990	5000	5010	5020	5030	5040	CTCCTGCCCT CTCACCCAT ACAGGGGACC CTGTCACTAC CTCTTGTCTC TCTCGTCACT GAATCCCTGA CTCCAGAACG CACTCAGCTC																					
L	L	A	S	H	A	I	Q	G	T	L	S	Y	P	L	V	S	V	V	S	E	S	L	T	P	E	R	T	Q	L	>
5050	5060	5070	5080	5090	5100	5110	5120	5130	CTCTATCTCC TTGCTGTGCTC TCTTGTCACT ATTCTGTAA TTATTCTGCT GGGGTAATC ATGCCAAAC GAAAGCTAA GCATGGCTC																					
L	Y	L	A	V	A	V	V	I	I	L	F	I	I	L	L	G	V	I	M	A	K	R	K	R	K	H	G	S	>	
5140	5150	5160	5170	5180	5190	5200	5210	5220	CTCTGGCTGC CTGAACGTTT CACTCTTCGC CGAGATGCAA GCAATCACAA GCGTCGTGAG CCAGTGGAC AGCATGCTGT GGGCTGAAA																					
L	W	L	P	E	G	F	T	L	R	R	D	A	S	N	H	K	R	R	E	P	V	G	Q	D	A	V	G	L	K	>
5230	5240	5250	5260	5270	5280	5290	5300	5310	AATCTCTCAG TGCAACTCTC AGAACGTAAC CTAATTGCTA CTGGAACAAG TGAACACTGG CTCCATGATC AAGGGCCCCA GCCAAAGAAA																					
N	L	S	V	Q	V	S	E	A	N	L	I	G	T	G	T	S	E	H	W	V	D	D	E	G	P	Q	P	K	K	>
5320	5330	5340	5350	5360	5370	5380	5390	5400	GTAAAGGCTG AACATGAGGC CTTACTCTCA GAAGAACATC ACCCCATTGA TCGACGGCCA TGGACACACC AGCACCTGA AGCTCCAGAC																					
V	K	A	E	D	E	A	L	L	S	E	E	D	D	P	I	D	R	R	P	W	T	Q	Q	H	L	E	A	A	D	>
5410	5420	5430	5440	5450	5460	5470	5480	5490	ATCCGTACCA CACCATCGCT GGCTCTCACC CCTCCTCAGG CAGAGCACCA CCTGGATGTC TTACATCTGA ATCTCCGTGC CCCAGATGGC																					
I	R	R	T	P	S	L	A	L	T	P	P	Q	A	E	Q	E	V	D	V	L	D	V	N	V	R	G	P	D	G	>
5500	5510	5520	5530	5540	5550	5560	5570	5580	TGCACCCCAT TGATGTTGGC TTCTCTCCGA GGAGCCAGCT CACATTGAG TGATGAAGAT GAACATCCAG AGGACTCTTC TGCTAACATC																					
C	T	P	L	M	L	A	S	L	R	G	G	S	S	D	L	S	D	E	D	E	A	E	D	S	S	A	N	I	>	
5590	5600	5610	5620	5630	5640	5650	5660	5670	ATCACAGACT TGGTCTACCA GGCTCCACC CTCCACGGCC AGACAGACCG GACTGGTAC ATGGCCCTGC ACCTTGCAGC CGCGTACTCA																					
I	T	D	L	V	Y	Q	G	A	S	L	Q	A	Q	T	D	R	T	G	E	M	A	L	H	L	A	A	R	Y	S	>

FIG.17G

5680	5690	5700	5710	5720	5730	5740	5750	5760
CGGGCTGATC CTGCCAAGCC TCTCCTGGAT GCACGGTCCAG ATCCCAATGC CCAGGACAAC ATGGGGCCCT GTCCACTCCA TCCTGGCAGTC R A D A A K R L L D A G A D A N A Q D N M G R C P L H A A V>								
5770	5780	5790	5800	5810	5820	5830	5840	5850
GCACCTGATC CCCAAGGTCT CTTCCAGATT CTGATTGGCA ACCGAGTAAC TGATCTAGAT GCCAGGATGA ATGATCCTAC TACACCCCTG A A D A Q G V F Q I L I R N R V T D L D A R M N D G T T P L>								
5860	5870	5880	5890	5900	5910	5920	5930	5940
ATCCTGCCCTG CCGGCTGGC TCTGGAGGG AATGGTGGCAG AACTGATCAA CTGCCAAGCC GATGTGAATC CAGTGGATGA CCATGGAAAA I L A A R L A V E G M V A E L I N C Q A D V N A V D D H G K>								
5950	5960	5970	5980	5990	6000	6010	6020	6030
TCTGCTCTTC ACTGGGCAGC TCTGTCAAT AATGTGGAGC CAACTCTTT CTGTTGAAA AATGGGGCC A C C G A G A C A T C G A C C G A C A A C S A L H W A A A V N N V E A T L L L L K N G A N R D M Q D N>								
6040	6050	6060	6070	6080	6090	6100	6110	6120
AACGAACAGA CACCTCTGTT TCTTGTGCC CGGGAGGGGA CCTATGAAGC AGCCAAGATC CTGTTAGACC ATTTGCCAA TCGAGACATC K E E T P L F L A A R E G S Y E A A K I L L D H F A N R D I>								
6130	6140	6150	6160	6170	6180	6190	6200	6210
ACAGACCATA TCGATCCTCT TCCCCGGAT GTGGCTGGG ATCGCATCCA CCATGACATT GTGGCCCTTC TGGATGAATA CAATGTGACC T D H M D R L P R D V A R D R M H H D I V R L L D E Y N V T>								
6220	6230	6240	6250	6260	6270	6280	6290	6300
CCAAGCCCTC CACCCACCT CTTGACTTCT GCTCTCTCAC CTGTCATCTG TGGGCCAAC AGATCTTCC TCAGCTGAA GCACACCCCA P S P P G T V L T S A L S P V I C G P N R S F L S L K H T P>								
6310	6320	6340	6350	6360	6370	6380	6390	6400
ATGGGCAACA AGTCTAGACG GCCCAGTCCC AAGACTACCA TCCCTACTAG CCTCCCTAAC CTTGCCAAGC AGGCAAGGA TCCCAAGGCT M G K K S R R P S A K S T M P T S L P N L A K E A K D A K G>								
6400	6410	6420	6430	6440	6450	6460	6470	6480
AGTACCGAGGA AGAACTCTCT GACTGAGAAC GTCCAACGT CTGAGACTTC ACTAACTTA TCCCTGTTC ATTCCCTAGA ATCTCCTCAC S R R K K S L S E K V Q L S E S S V T L S P V D S L E S P H>								

FIG.17H

6490	6500	6510	6520	6530	6540	6550	6560	6570
ACCTATGTTT	CCGACACCA	ATCCCTCCA	ATGATTACAT	CCCCTGGAT	CTTACAGGCC	TCACCCAACC	CTATCTGCC	CACTGGCCCC
T Y V	S D T T	S S P	M I T	S P G I L	Q A S P N	P M L A	T A A >	
6580	6590	6600	6610	6620	6630	6640	6650	6660
CCTCCCTGCC	CACTCCATGC	CCACCATCCA	CTATCTTTT	CTAACCTICA	TCAAATGCAG	CTTTGCCAC	ATGGGGCCAG	CACTGTCCTT
P P A	P V H A	Q H A L	S F	S N L H	E M Q	P L A H	G A S T	V L >
6670	6680	6690	6700	6710	6720	6730	6740	6750
CCCTCAGTCA	CCCAGTGTCT	ATCCCACCA	CACATTGTCT	CTCCAGGCAG	TGGCACTGCT	CGAACCTTCA	GTAGGCTCCA	TCCAGTCCCA
P S V	S Q L L	S H H	H I V	S P G S	G S A	G S L	S R L H	P V P >
6760	6770	6780	6790	6800	6810	6820	6830	6840
GTCCCACCA	ATTGGATCAA	CCGCATGGAC	GTGAATGAGA	CCCAGTACAA	TGAGATGTTT	GGTATGGTCC	TGGCTCCAGC	TGAGGGCACC
V P A	D W M N	R M E	V N E	T Q Y N	E M F	G M V	L A P A	E G T >
6850	6860	6870	6880	6890	6900	6910	6920	6930
CATCCTGCCA	TAGCTCCCCA	GAGCAGGCCA	CCTGAAGGCA	AGCACATAAC	CACCCCTCGG	GACCCCTTGC	CCCCATTGT	GACTTCCAG
H P G	I A P Q	S R P	P E G	K H I T	T P R	E P L	P P I V	T F Q >
6940	6950	6960	6970	6980	6990	7000	7010	7020
CTCATCCCTA	AAGCCAGTAT	TGCCCAACCA	GGGGGGGCTC	CCCAGCTCA	GTCCACCTGC	CCTCCAGCTG	TTGGGGGCC	CCTGCCACC
L I P	K G S I	A Q P	A G A	P Q P Q	S T C	P P A	V A G P	L P T >
7030	7040	7050	7060	7070	7080	7090	7100	7110
ATCTACCAGA	TTCCAGAAAT	GGCCCGTTG	CCCAGTGTGC	CTTCCCCAC	TGCCATGATC	CCCCACCCAGG	ACCGGCAGGT	ACCTCAGACC
M Y Q	I P E M	A R L	P S V	A F P T	A M M	P Q Q	D G Q V	A Q T >
7120	7130	7140	7150	7160	7170	7180	7190	7200
ATTCTCCCA	CCTATCATCC	TTTCCCAGCC	TCTGTGGCA	AGTACCCAC	ACCCCTICA	CACCAAGCTT	ATGCTTCCTC	AAATGCTGCT
I L P	A Y H P	F P A	S V G	K Y P T	P P S	Q H S	Y A S S	N A A >
7210	7220	7230	7240	7250	7260	7270	7280	7290
GAGCGAACAC	CCAGTCACAG	TGGTCACCTC	CACCGTGACC	ATCCCTACCT	GACACCATCC	CCAGAGTCIC	CTGACCACTG	GTCAAGTICA
E R T	P S H S	G H L	Q G E	H P Y L	T P S	P E S	P D Q W	S S S >

FIG.171

7300	7310	7320	7330	7340	7350	7360	7370	7380
TCACCCCACT	CTGCTCTGA	CTGCTCAGAT	CTGACCACCA	GCCCTACCCC	TGGGGCTCT	GGAGGAGCTC	ACCGGGGACC	TGGGACACAC
S P H S A S D	W S D	V T T	S P T P	I G G A	G G G	Q R C P	G T H D	
7390	7400	7410	7420	7430	7440	7450	7460	7470
ATCTCTGACC	CACCACACAA	CAACATCCAG	GTITATCGCT	GAGAGACTCC	ACCTCCACTG	TACAGACATA	ACTGACTTTT	GTAAATGCTG
M S E	P P H N	N M Q	V Y A>					
7480	7490	7500	7510	7520	7530	7540	7550	7560
CTGACCAACA	AATGAACCTC	ATCCCCGAGA	GAAATGAACA	AATCTCTGGA	GCCACCTTCT	ACAGGTAGGA	AAGAGAACAT	GTCTTATTTC
7570	7580	7590	7600	7610	7620	7630	7640	7650
AGATAATCCA	ACAGAACCAA	TTCTCTAGTT	TCACTGGTA	TCTGCAAGGC	TTATTGATTA	TTCTAATCTA	ATAAGACAAG	TTTGTGAAA
7660	7670	7680	7690	7700	7710	7720	7730	7740
TCCAACATGA	ATACAACCC	TGGTCCATG	TTTACTCTCT	TCTATTTGGA	CAATAACATG	GATGCTTATT	GAAGCCCAGA	CATTCTGCA
7750	7760	7770	7780	7790	7800	7810	7820	7830
GCTTGGACTC	CATTTAACCC	CCTGCAGGCT	TCTGCCATAT	CCATGAGAAC	ATTCTACACT	AGCCTCCTGT	TGGCAATTAT	GCCCTGGAAAT
7840	7850	7860	7870	7880	7890	7900	7910	7920
TCTGCCCTGAA	TTGACCTACG	CATCTCCTCC	TCCTTGGACA	TTCATTTGTC	TTCATTTGCT	GCTTTGCTT	TTGACCTCT	CCCTGATTGT
7930	7940	7950	7960	7970	7980	7990	8000	8010
ACCCCTACCA	CCATGTTATA	GGGCAAGACC	TTTGTCTTT	TGATCATTCT	GGCCCCATGAA	AGCAACTTTC	GTCTCCTTTC	CCCTCCTGTC
8020	8030	8040	8050	8060	8070	8080	8090	8100
TTCCCCGGTAT	CCCTTGGACT	CTCACAAAGT	TTACTTTGGT	ATCGTTCTCA	GCACAAACCT	TTCAACTATC	TTCTTTCTTT	GGAAAATGCA
8110	8120	8130	8140	8150	8160	8170	8180	8190
CATACTGTAT	TGTCTCTCC	TGGCATATATC	ATTCTGGAC	AGAGAAGGGG	AGAAGAACATC	TTTTCTCAA	CAAATTTGCG	GGGCAGGAGA
8200	8210	8220	8230	8240	8250	8260	8270	8280
TCCCTTCAAG	AGGCTGGACC	TTAATTTTC	TTCTCTGTGT	CCAGGTCTTC	ATATAAACTT	TACCAAGGAAC	AAGGCTGTCA	CTTCTTGTCT

FIG.17J

67

8290	8300	8310	8320	8330	8340	8350	8360	8370
TTTCCTGTA TGGCCCTGGT CAGTCTAAAG TTTTATCCTT GATAGTCTAG TTACTATGAC CCTCCCCACT TTTTAAAAC CAGAAAAAGG								
8380	8390	8400	8410	8420	8430	8440	8450	8460
TTTGGAAATCT TGGAATGACC AACAGACAAC TAACTCGTC CAAGACCCAG TTACCCACCC ACAGGTCCCC CTACTTCCTG CCAAGCATT								
8470	8480	8490	8500	8510	8520	8530	8540	8550
CATTGACTCC CTCTATCGAA CACATTTCTC CCAGATCTGA GCATTCTAGG CCTCTTCAC TCACTCACCC AGCATATGAA ACTAGCTTA								
8560	8570	8580	8590	8600	8610	8620	8630	8640
ACTCTTGAGC CTTCCCTTC ATATCCACAG AAGACACTGT CTCAAATGTT GTACCCCTGC CATTAGGAC TGAACTTCC TTAGCCCAAG								
8650	8660	8670	8680	8690	8700	8710	8720	8730
GGACCCAGTG ACAGTTGCTC TCCGTTCTC AGATGATCAC TCTCTACTGA TTATCTTCCT CCTTAAAGGC CTGCTCACCA ATCTTCTT								
8740	8750	8760	8770	8780	8790	8800	8810	8820
CACACCGTGT GGTCCGTGTT ACTGGTATAC CCACTATGTT CTCACTGAAG ACATGGACTT TATATGTTCA ACTCCAGGAA TTGGAAAGTT								
8830	8840	8850	8860	8870	8880	8890	8900	8910
GGACTTGT TCTATGATCC AAAACAGCCC TATAAGAACG TTGGAAAAGG AGGAACATA TAGCAGCCTT TGCTATTTTC TGCTACCATT								
8920	8930	8940	8950	8960	8970	8980	8990	9000
TCTTTCCCTC TGAACCGGCC ATGACATTCC CTTGGCAAC TAACGTAGAA ACTCAACACA ACATTTCCCT TTCCCTAGAGT CACCTTTAG								
9010	9020	9030	9040	9050	9060	9070	9080	9090
ATCATAATCC ACAACTATAC ACTTGCTCAT TCTTCAGACT GATTCCCCCT CACCTGAATC CACTCTCTGT ATTCACTGCTC TTGGCAATT								
9100	9110	9120	9130	9140	9150	9160	9170	9180
CTTTCACCTT CTTTAACGG CAGAACCAATT TTACTTAATT GTAGATAAAG AATACTTTTC TTCCCTCTCT CCTTGGGCCA CTTAATAATT								
9190	9200	9210	9220	9230	9240	9250	9260	9270
GGTCCATGGC TACACTGCAA CTTCCGTCCA GTGCTGTGAT CCCCATGACA CCTGCAAAT AACTCTGCC TGGCCATTCTT GTAGATATTA								

FIG.17K

9280 9290 9300 9310 9320 9330 9340 9350 9360
ACACGTGAAT TCCCAGCTCTT TTTGCTTCA ATGACAGTC TCATTCCTTC TATGGCTGCA ACTATCCATC ACTGCTTCCC ACTTACCTGA
9370 9380 9390 9400 9410 9420 9430 9440 9450
TTTGTCTGTC CCTGGCCCCA TATGGAAACC CTGGCTGTCCTT GTTGGCATAAA TACTTACAA ATGGTTTTTT CACTCCTATC CAAATTTATT
9460 9470 9480 9490 9500 9510 9520 9530 9540
GAACCAACAA AAATAATTAC TTCTGCCCTG AGATAAGCAG ATTAAGTTTG TTCATTCTCT GCTTTATTCT CTCCATGTGG CAACATTCTG
9550 9560 9570 9580 9590 9600 9610 9620 9630
TCAGCCTCTT TCATACTGTC CAAACATTAA ATCATTCTAA ATGGTCACTC TCTGCCCTTG GACCCATTAA TTATTCACAG ATGGGGAGAA
9640 9650 9660 9670 9680 9690 9700 9710 9720
CCTATCTCCA TCCACCCCTCA CCATCCTCTG TCCAGCACAC ACAGTCCAGG GAGCCAGTGG CCATGGCGAT GACTTCTTC CCCTGGGAAT
TCC

FIG.17L